OPTIMIZING CHOCOLATE CHIP COOKIE RECIPES FOR RATINGS USING MACHINE LEARNING AND DEEP VECTOR-TO-SEQUENCE RNN MODELS

Mackenzie O'Brien

Dr. Pablo Rivas, Faculty Advisor

Spring 2019

Submitted to the Honors Council in partial fulfillment of the requirements for the degree of Honors in Liberal Arts

Abstract

This research studies machine learning algorithms in the process of producing new recipes for cookies. The objective of these experiments was to generate a cookie recipe that was optimized for critic ratings by using multiple types of neural and non-neural networks to predict and create chocolate chip cookie recipes based on the input of over 250 human-made recipes and instructions. There were 138 different parameters inputted, including Rating, Calories, and 136 different ingredients such as sugar, flour, and egg. To get the instructions, we created a vector-to-sequence algorithm that takes the input of a recipe ingredient vector and uses the instructions from the 250 man-made recipes to make predictions about the sequence of instructions in the output. Results suggest that some cookies baked using machine learning are likeable by human subjects.

1 Introduction

In 2017, Google published a blog post about their experiments in designing the best possible chocolate chip cookies using a given set of ingredients [Golovin, 2017]. In searching for a topic for my honors thesis, I came across this article and was intrigued by the idea of making a chocolate chip cookie recipe from machine learning and not trial and error. As a computer science major, most problems that are approached are super technical and not thought of as "fun" projects. My motivation for researching the optimization of cookies stems from wanting a topic that strayed from the typical projects in the field of computer science and from having a love for baking.

Goal: To Make a Highly Rated Chocolate Chip Cookie

The objective of this research is to generate cookie recipes that are optimized for critic ratings. This is done by using multiple types of neural and non-neural networks to predict and create chocolate chip cookie recipes based on the input of over 250 human-made recipes and instructions. There were 138 different parameters inputted, including Rating, Calories, and 136 different ingredients such as sugar, flour, and egg.

Goal: To Produce Usable and Unique Baking Instructions

To get the instructions, we created a vector-to-sequence algorithm that takes the input of a recipe ingredient vector to use as an attention mechanism and uses the instructions from the 250 man-made recipes to make predictions about the sequence of instructions in the output.

Related Work

Naik's study uses `allrecipes.com` for data collection. They employ a generative model based on Bayesian pairwise probabilities calculated from collected recipes. They then use the input taken in and the ingredient pairings to generate ingredients and instructions. This inspired our work to use the same website to collect data and to also use generative models. Similarly to [Naik and Polamreddi, 2015], Google^{TM} attempted to use AI to find the perfect chocolate chip cookie, they also employed Bayesian Optimization [Clifford, 2018]

The use of machine learning for baking/cooking is relatively unexplored, but there are some articles, such as *Bayesian Optimization for a Better Dessert* [Solnik et al., 2017] and *Cuisine Classification and Recipe Generation* [Naik and Polamreddi, 2015] that have addressed the subject matter. For example, Kochanski's Bayesian Optimization study applied Bayesian Optimization in an effort to optimize chocolate chip cookies and was comprised of a mixed system of human chefs, raters, and a machine optimizer in

144 experiments across the country. They used a Vizier black-box optimization tool for new recipes using a Bayesian Optimization bandit algorithm based on Gaussian Process model. When conducting exploratory research, we were unable to find another study that employed vector-to-sequence algorithms in the manner that ours does. [Solnik et al., 2017]

The rest of the paper is organized as follows: Section 2 examines our methodology of machine learning by delving into separate subsections for ingredients and instructions, Section 3 addresses our process of baking and serving the cookies and our survey design, Section 4 begins our analysis of results both for ingredients and instructions by looking at performance metrics, and Section 5 details our lessons learned, followed by our conclusion in Section 6. In Section 8 our supplemental materials are displayed in detail, followed by references.

2 Machine Learning Methodology

In our research, the production of new recipes were split into two stages: computation of ingredients and the generation of instructions. First, we used different algorithms to return ingredient vectors which were then fed into a LSTM recurrent neural network to generate instructions for that specific vector of ingredients. This last process is known as vector-to-sequence modeling.

2.1 Ingredient Selection

This experiment involved testing 9 different algorithms for generating ingredient vectors: Deep Neural Networks, Extremely Randomized trees,Gradient Boosting, Linear Regression, Neural Networks, Normalized Neural Networks, Wide Neural Networks,Random Forest, Support Vector Machine (SVM). Our target variable $\mathbf{y}$ was the 'rating' of a cookie.

By using `allrecipes.com` as our dataset, we were able to target our predicted rankings off of the ratings included with each recipe in our training data set. Each algorithm generated new cookie recipes by selecting a previously existing value for each ingredient column and analyzing how the combination compared to similar recipes by calculating uniqueness and simplicity metrics. We defined simplicity as follows:

$$Simplicity = ||y - y^*|| + ||\mathbf{w}|| \tag{1}$$

where y is the true ranking, y^* is the predicted ranking, and $\mathbf{w}$ is the vector of weights associated with the regression problem given by $\mathbf{w}^T\mathbf{x} = y$. Then we define the uniqueness metric as follows:

$$Uniqueness = \min_{i=1,\ldots,N} ||\mathbf{x}_i - \mathbf{x}^*|| \tag{2}$$

where $\mathbf{x}_i$ is the i-th sample vector from the ingredients data set, $\mathbf{x}^*$ is the new/proposed set of ingredients and N is the size of the data set.

Neural Networks

Neural Networks are known as artificial networks and may have multiple layers between the input and output layers; they are said to approximate any non-linear function [Hajian and Styles, 2018]. The layers combine into a complex manipulation of data that is possible through powerful computing that attempts to replicate the human brain. Each mathematical manipulation of the data is considered a layer, and complex neural networks have many layers. In our research, we trained 4 types of neural networks as described below.

Deep Neural Networks: Deep Neural Nets (DNNs) for regression are a type of neural network with multiple layers between the input and output layers. Each mathematical manipulation of the data is considered a layer and a level of data abstraction; complex DNNs have many layers [Glorot and Bengio, 2010]. Figure 1 display s an example of a DNN with 6 layers.

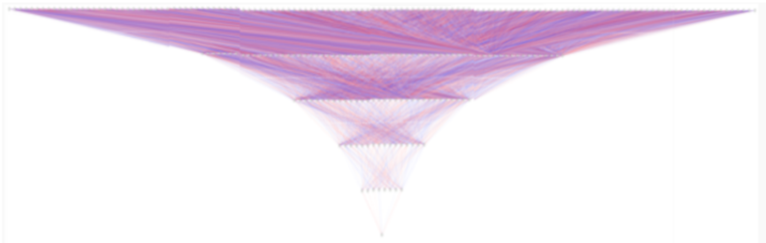


Figure 1: Deep Neural Network

Wide Neural Network: A wide neural network has less layers than a deep neural network, but usually has a larger number of neurons per layer, thus becoming "wider" not "deeper". Figure 2 shows an example of a three layer network with a large number of neurons in its layers.

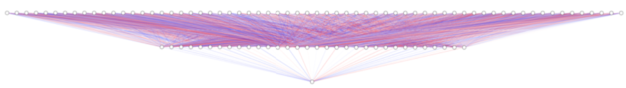


Figure 2: Wide and Shallow Neural Network

Shallow Neural Network: A shallow neural network is a basic neural network that uses fewer hidden layers in its computation, often times only using one hidden layer. While not as complex, computation time is much faster than a deep neural network. An example of shallow neural network in shown in Figure 2.

Normalized Neural Network: In this research, we refer to a normalized neural network as simply a shallow neural network that has had its input scaled using a z-score approach. This means that the values have been scaled to have zero-mean and unit variance using the following formula:

$$\mathbf{x}_z = \frac{\mathbf{x} - \mu}{\sigma} \tag{3}$$

where μ is a vector of sample means over the columns in $\mathbf{x}$ and σ is a vector of standard deviations for each column in $\mathbf{x}$.

The process of scaling a neural network input is known to have boosted performance for some shallow networks [Jolai and Ghanbari, 2010].

Other Machine Learning Algorithms

In our experiments we also used other non-neural network-based algorithms as described next.

Random Forests: The *Random Forest* algorithm is an algorithm that uses a "forest" of decision trees and combines the results in "bagging" to generate an overall result/prediction [Donges and Donges, 2018]. The model uses two key concepts: random sampling of training data points when building trees and random subsets of features considered when splitting nodes. Figure 3 shows an example of the shape of a decision tree and the branches made by the values in each feature.

Extremely Randomized Trees: The Extremely Randomized Trees algorithm is a variant of the Random Forest regression algorithm, but differs in the fact that at each step of the extreme tree, the entire sample is used and a decision boundary is picked at random, rather than the locally best one out of a small sampling of the set [Donges and Donges, 2018]. The extremely randomized trees algorithm was chosen because it encompasses the ideas behind both itself and the Random Forest algorithm, which were very similar in their results and processes.

Gradient Boosting: The *gradient boosting* algorithm applies the concept of modifying a weak learner to become better at identifying good outcomes more efficiently [Brownlee, 2016]. The model, like extremely randomized trees, uses an ensemble of weak prediction models, typically decision trees, to

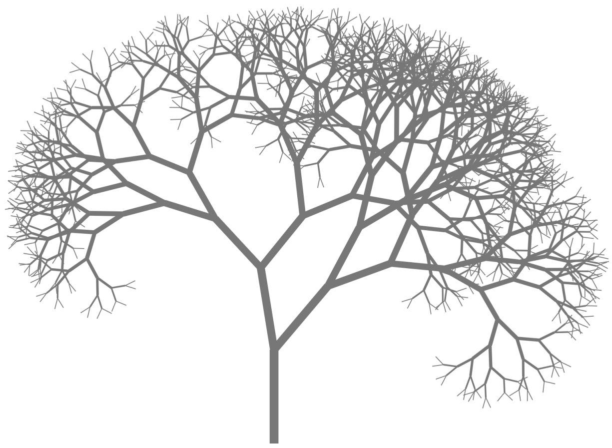


Figure 3: Decision Tree Forest

make predictions. However, unlike extremely randomized trees, the gradient boosting algorithm also attempts to minimize errors through a loss function.

Support Vector Machines: The objective of support vector machines is to find a hyper-plane in an M-dimensional space (M = the number of features) that distinctly classifies the data points [Rivas-Perea et al., 2013]. When generating random new recipes, an SVM classifies the new recipes according to the classification hyper-planes created during training.

Linear Regression: Linear Regression is used to predict a value for a target variable based on values of a predictor variable and regression coefficient. Linear Regression relies heavily on correlation among variables independently and as sets of variables [Montgomery et al., 2012]. In our research, multiple linear regression was used to include multiple predictors and one target. This algorithm was used to predict ranking with the ingredients as independent variables and rankings as the dependent variable. A generated ingredients vector contains the the names of the each ingredient and the measurement of that ingredient; an example of an ingredients vector is as follows:

sugar(c)	avocado	coconut(c)	flour(c)	vanilla(tsp)	choc. chips(c)
.75	1	.3	3	1	3
1	0	0	2	0	2

2.2 Instructions Production

After training the above described algorithms to output recipes with high predicted rank, low simplicity, and high uniqueness, our next goal was to train our instruction algorithm to generate recipes that included all of the steps for baking based on an ingredients vector.

The algorithm uses a Long Short-Term Memory (LSTM) recurrent neural network, as illustrated in Figure 5, which incorporates each generated recipe as an attention mechanism [Wang et al., 2016]. By creating a modified merge model with the attention mechanism for ingredients, we attempted to train the model to better predict words in the instructions.

The text input for each recurrent layer includes the ingredients vector, the previously outputted text, and the target output. For example, if the network had the ingredient vector and the words "Preheat the", the target word associated with those inputs would be "oven".

The model was trained to have 150 recurrent layers of the LSTM, each with 932 outputs with soft-max activation. The length the sequence of 150 words was chosen by analyzing our instructions data set in order to achieve the inclusion of full length recipes for about 98% of the data set: as shown in Figure 4. This results in recipes of 150 words, one word of output per iteration. The soft-max activation works by selecting the word with the highest probability as the correct output. To pass the previous word into the next LSTM iteration, we used a word embedding process that translated the word back into vector form. The cross entropy loss was calculated for each sequence produced, which analyzes how accurate a sequence is compared to instructions in our data set.

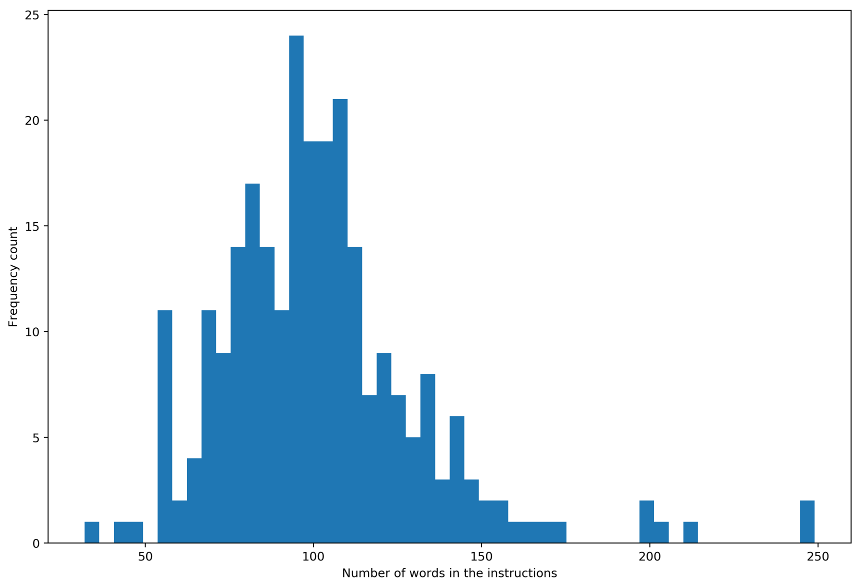


Figure 4: Histogram of instruction lengths.

The cross entropy loss is defined as:

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{c\in C}\mathbf{d}^*_{cn}\ln\mathbf{d}_{cn}+(1-\mathbf{d}^*_{cn})\ln(1-\mathbf{d}_{cn}) \tag{4}$$

where $\mathbf{d}_n \in \mathbb{R}^{932}$ is the true probability of the n-th sample belonging to a specific word, c, in the dictionary of words for directions. The size of such dictionary is 932.

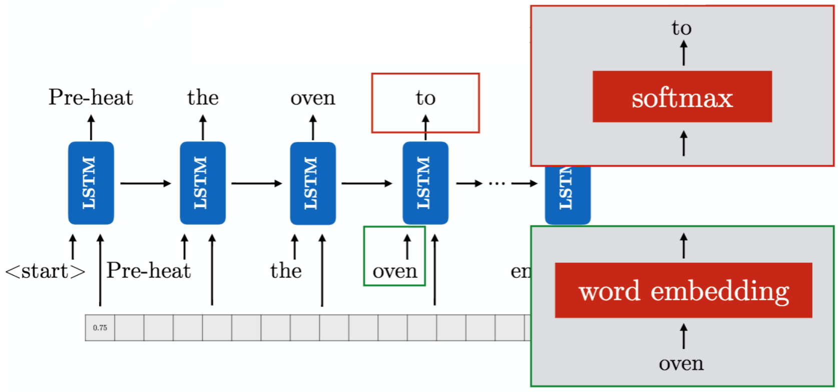


Figure 5: Our merge model with an ingredients attention mechanism.

3 Baking, Serving, and Surveying Methodology

After designing the algorithms, we proposed 5 different taste test experiments, each time testing a cookie designed with our algorithms against our control cookie, the well-known Nestle® Toll House® chocolate chip cookie [Times, 2018]. The following paragraphs explain the baking and serving protocols.

3.1 Baking and Serving Protocol

Each taste experiment consisted of 40 cookies each for the control and the machine learning recipes, and the cookies were baked the day before each experiment to ensure freshness and quality remained the same for each test. Our procedure for serving the cookies is as follows:

1. We explained to the participant the risks associated with this experiment and made available to the participant a printed copy of the document entitled "Waiver and Release to Medical Attention and Grant of Rights" for further reading and answered any questions before proceeding. See the supplemental materials section for a copy of the documents.

2. We then provided each participant two cookies: the experimental cookie and the control cookie in a single-blind taste test. Samples were placed in small bags labeled Cookie 1 or Cookie 2. Every bag contained labels with QR codes and links to an online survey for each cookie, and the table had a napkins and cups of water available for each participant.

3. Participants were discouraged from talking to one another during the tasting event and were not able to see how other participants are scoring each sample.

4. We asked each participant to test one cookie by first recording their score for appearance, then aroma, then taste, and finally texture. Note that texture pertains to how the food feels in your mouth. For example: crunchy, chewy, juicy, soggy, creamy, and so on. See supplemental materials for a copy of the survey.

5. After testing the sample, we provided water for the participants to cleanse their palate. We then asked the participants to repeat steps three and four for the second sample cookie.

3.2 Survey Design

For each cookie that a participant tasted, we asked them to complete a survey giving their consent to use their information and questions about different attributes of the cookie and their overall satisfaction with the cookie. Survey questions included:

- Appearance on a scale of 1 (Unfit for consumption) to 5 (Excellent)
- Aroma on a scale of 1 (Unfit for consumption) to 5 (Excellent)
- Taste on a scale of 1 (Unfit for consumption) to 5 (Excellent)
- Texture:
  - Crunchy
  - Chewy
  - Gooey
  - Juicy
  - Soggy
  - Creamy
  - Other
- Overall Satisfaction on a scale of 1 (Hated It) to 10 (Loved It)

In our research, we are most interested in the results of overall satisfaction with the cookie, as the goal is to produce a recipe with optimized ratings. Our definition optimized rating is a cookie that receives high average ratings for overall satisfaction. A sample survey is shown in the supplemental materials section.

4 Analysis

In this section we will discuss the process of analysis of results in their different areas. We begin with the models trained over ingredients, and how we assessed their quality; and then models to learn to produce instructions for baking the ingredients and performance metrics during and after training. Finally, we go in depth in the analysis of the survey results.

4.1 Ingredients Selection

To analyze the algorithms that generated ingredient vectors, mean-squared error loss and the Coefficient of Determination were calculated for each algorithm, and the best algorithms were chosen to pick recipes from. Based on the mentioned metrics from each, shown in Table 1, we narrowed our focus to the following three algorithms: Deep Learning, Gradient Boosting, and Extremely Randomized Trees. We chose the Deep Learning and Extreme Trees Algorithms because the mean squared error loss was low and the coefficient of determination,R^2, was high, while the Gradient Boosting Algorithm was chosen to represent the other side of the spectrum with high loss and R^2. Mean squared error loss is defined as follows:

$$\frac{1}{N}\sum_{i=1}^{N}(y_i - y^*)^2 \tag{5}$$

and the coefficient of determination is defined as:

$$R^2 = 1 - \frac{u}{v} \tag{6}$$

where u is the residual sum of squares $\sum_{i=1}^{N}(y_i - y^*)^2$ and v is the total sum of squares $\sum_{i=1}^{N}(y_i - \bar{y})^2$ Here, $\bar{y}$ indicates the mean of y. The appendix contains the actual recipes selected for baking.

As shown in Table 2, five recipes were then selected from the three chosen algorithms to test. We chose to do two of each from the better scoring algorithms and one from Gradient Boosting. When choosing these recipes, we sorted first by predicted rank high to low, then by simplicity low to high and uniqueness high to low. After sorting, we chose one of the top few recipes for each.

Table 1: Mean-Square Error and Coefficient of Determination.

Algorithm	**MSE**	R^2
Deep Neural Network	0.0231	0.9267
Extremely Randomized Trees*	**0.0001**	**1.0**
Gradient Boosting	0.1111	0.6481
Linear Regression	0.1056	0.6656
Normalized Neural Network	**0.0219**	**0.9305**
Random Forests	0.0763	0.7584
Wide Neural Network	**0.0204**	**0.9354**
Shallow Neural Network	0.0429	0.8634
SVM	0.1873	0.4068

*ERTs keep copies of the data and usually yields perfect correlation scores.

Table 2: Cookie Recipes and Their Uniqueness and Simplicity

Batch	**Algorithm**	**Simple**	**Unique**	**P. Rank**
A	Extremely Randomized Trees	14.1	5.4	5.0
B	Gradient Boosting	20.3	5.8	5.1
C	Extremely Randomized Trees	**4.1**	2.3	5.0
D	Deep Learning	11.6	**6.1**	9.5
E	Deep Learning	17.3	5.0	9.4

4.2 Instructions Productions

As the vector-to-sequence algorithm is previously untested in other research, the end results leave something to desire in terms of inclusion of all ingredients and actual usability. However, it is an accomplishment to have gotten a working algorithm that takes in an ingredient vector and outputs a semi-usable recipe. Improvements include ensuring that the instructions contain all ingredients in the vector that contain non-zero values and eliminating repeating loops that the algorithm gets stuck on. Perhaps if we had a larger training data set, the algorithm could learn to better include ingredients from the attention mechanism vector.

In Figure 6, each run is shown with the corresponding loss score against each epoch. The length of each line indicates how many epochs training lasted until the Loss stopped decreasing. Figure 6 suggests that in most cases five epochs is enough for convergence. The goal and result of the model was to train it to reduce the cross entropy loss score. Figures 7 and 8 show the BLEU scores for each run. Figure 7 shows the BLEU score for each run and the corresponding epoch in which the score was calculated.

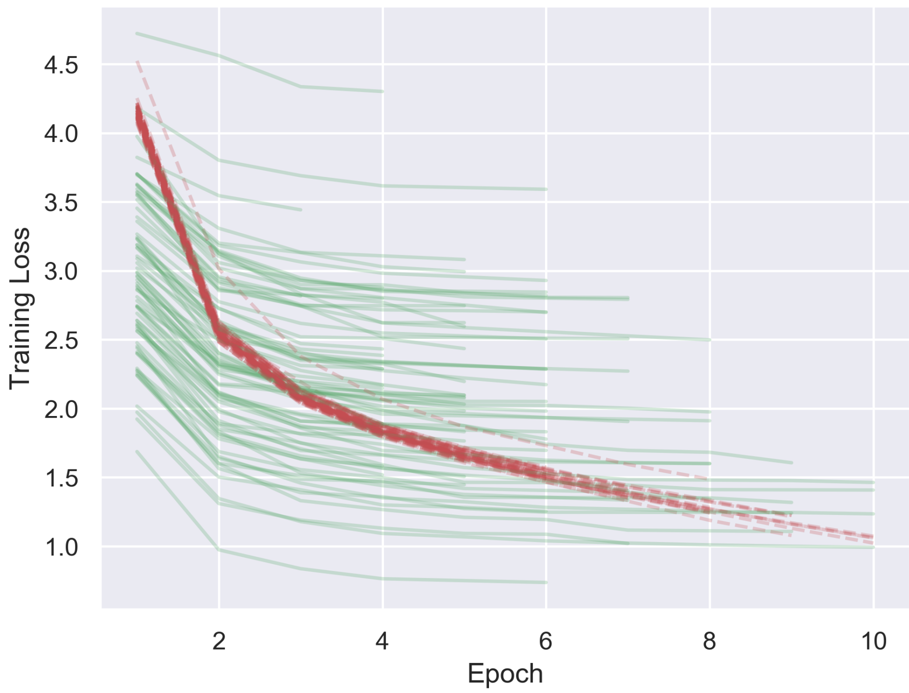


Figure 6: Loss Score Per Epoch

Figure 8 shows the range of scores for each epoch, showing that the best runs reached the lowest loss at epoch nine.

A BLEU score (bilingual evaluation understudy) is an algorithm for evaluating the quality of text which has been machine-translated from one natural language to another. Quality is considered to be the difference between a machine's output and that of a human. In other words, the closer a machine's output to that of a human's, the better it is. The output is always between 0 and 1, and the higher the score, the better the machine output, and the score is gathered by individually calculating segments (generally sentences) and averaging the results for an estimate on overall quality [Papineni et al., 2002]. .

The supplemental materials section shows the text generated as instructions for each recipe.

4.3 Survey analysis

As shown in Figures 9 and 10 none of the Machine Learning generated cookies scored better in terms of overall satisfaction. However, cookie B came the closest. Cookies C and D scored the worst, proving our hypotheses that a simple and not very unique cookie nor a non-simple but moderately unique cookie are not always the best choices. The Deep Neural Network

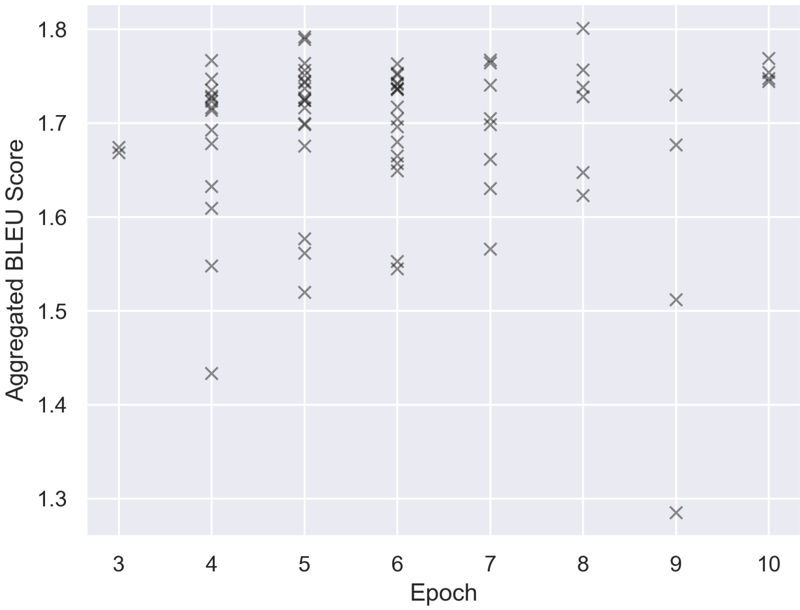


Figure 7: BLEU Scores Scatter Plot

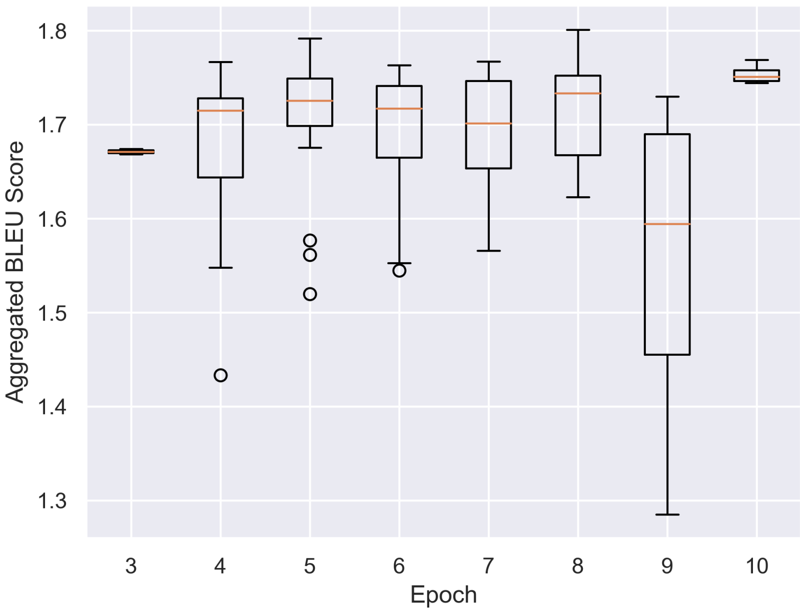


Figure 8: BLEU Scores Box Plot

cookies fared the worst as cookies D and E. Cookie E had to be modified to form a proper dough, as the generated recipe contained 0 dry ingredients;

surprising since the simplicity score suggested a complex cookie. Figure 9 shows the histogram of the responses to the rating of each cookie. The figure suggests that A,B, and E and rated higher than the rest. In Figure 10, the cookies that were rated most similarly to the control were batches A, B, and E, which is to be expected based on results from Figure 9. Figure 10 depicts the result of creating a histogram out of the difference in rating that a subject gives to our cookie compared to their rating of the control cookie; i.e. Δ ML-Control.

Cookie B, an Extremely Randomized Tree cookie was rated the highest in our experiments. This result was surprising because although it had the highest simplicity score (fewest ingredients) and the second highest uniqueness score, it was tied for last in terms of predicted ranking.

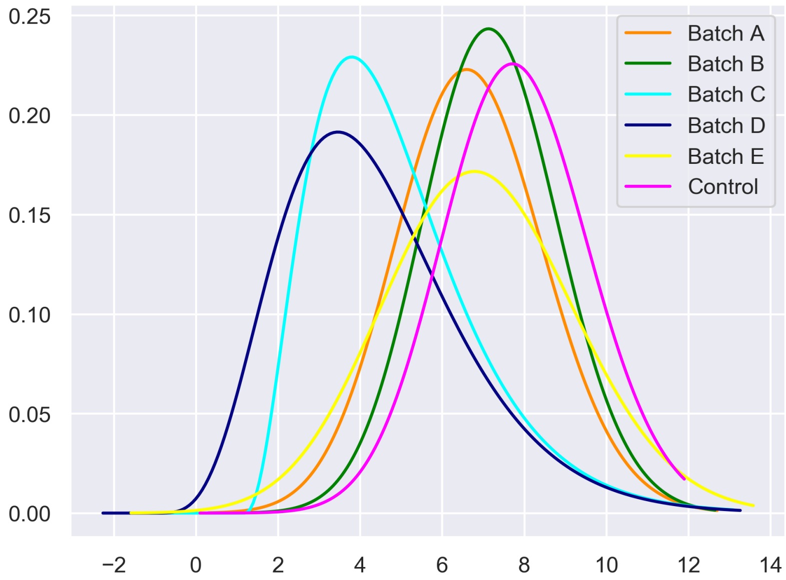


Figure 9: Overall Satisfaction

5 Lessons Learned

Our goal was to learn and illustrate the potential of machine learning in a real-world setting in a domain not typically thought of as appropriate for computer interaction. There is still much to be explored in terms of applying machine learning to areas outside of the worlds of finance and mathematical calculations. As we were doing our experimental tests in between and during classes at the college, it was sometime hard to get participation in a timely manner, and it would have been better in hindsight to get committed testers for all five taste tests or host the tasting at larger events on campus. Perhaps

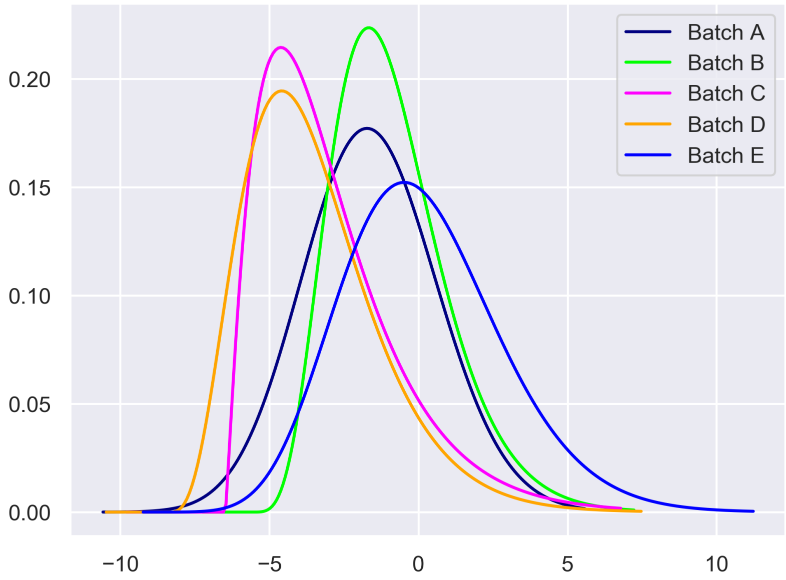


Figure 10: Δ ML - Control

if we had done all five cookies at one event, we could have gotten different results as well. Future work will include analysis of the other points of evaluation in the survey.

6 Conclusion

The results of our research have set a good base line for further research on this subject. After much analysis of nine algorithms, we have proven that good chocolate chip cookie recipes can be generated through machine learning and that this work can be applied to other recipes provided that a training set is created. The best recipe came from the Extremely Randomized Trees algorithm, and the decision tree algorithms did better in general, while the Deep Learning recipes proved to be the worst of the 5 chosen.

Our research on vector-to-sequence models has room for much expansion on the topic and is only just a start. The results from those experiments prove that our model is headed in the correct direction but needs further investigation to get unique recipes for inputted ingredient vectors.

Acknowledgments

We would like to thank Marist College as whole, the Computer Science Department, and the Honors Department for their continued support of this research, as well as the IRB for their review and our cookie taste testers for their participation in our experiments.

References

[Brownlee, 2016] Brownlee, J. (2016). A gentle introduction to the gradient boosting algorithm for machine learning. *Machine Learning Mastery. Nov*, 9.

[Clifford, 2018] Clifford, C. (2018). Alphabet billionaire eric schmidt: Google used a.i. to find the perfect chocolate chip cookie recipe.

[Donges and Donges, 2018] Donges, N. and Donges, N. (2018). The random forest algorithm.

[Glorot and Bengio, 2010] Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256.

[Golovin, 2017] Golovin, D. (2017). The makings of a smart cookie.

[Hajian and Styles, 2018] Hajian, A. and Styles, P. (2018). Artificial neural networks. In *Application of Soft Computing and Intelligent Methods in Geophysics*, pages 3–69. Springer.

[Jolai and Ghanbari, 2010] Jolai, F. and Ghanbari, A. (2010). Integrating data transformation techniques with hopfield neural networks for solving travelling salesman problem. *Expert Systems with Applications*, 37(7):5331–5335.

[Montgomery et al., 2012] Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to linear regression analysis*, volume 821. John Wiley & Sons.

[Naik and Polamreddi, 2015] Naik, J. and Polamreddi, V. (2015). Cuisine classification and recipe generation. *Online: http://cs229. stanford. edu/proj2015/233 report. pdf.*

[Papineni et al., 2002] Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.

[Rivas-Perea et al., 2013] Rivas-Perea, P., Cota-Ruiz, J., Chaparro, D. G., Venzor, J. A. P., Carreón, A. Q., and Rosiles, J. G. (2013). Support vector machines for regression: a succinct review of large-scale and linear programming formulations. *International Journal of Intelligence Science*, 3(01):5–14.

[Solnik et al., 2017] Solnik, B., Golovin, D., Kochanski, G., Karro, J. E., Moitra, S., and Sculley, D. (2017). Bayesian optimization for a better dessert.

[Times, 2018] Times, T. N. Y. (2018). Toll house chocolate chip cookies.

[Wang et al., 2016] Wang, Y., Huang, M., Zhao, L., et al. (2016). Attention-based lstm for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 606–615.

7 Supplemental Material

7.1 Final Recipes with Unedited Instruction Sequences

`Finished Cookies`

Figure 11: Cookie A

Figure 12: Cookie B

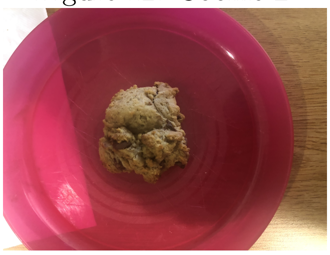


Figure 13: Cookie C

Figure 14: Cookie D

Figure 15: Cookie E

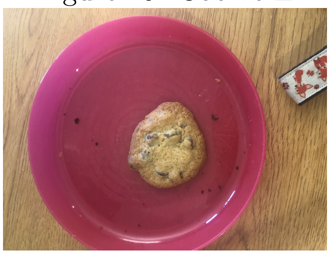


Figure 16: Control

Control Cookie: Nestle® Toll House® Cookie

Ingredients:

- 2 1/4 cups all-purpose flour
- 1 teaspoon baking soda
- 1 teaspoon salt
- 1 cup (2 sticks) butter, softened
- 3/4 cup granulated sugar
- 3/4 cup packed brown sugar
- 1 teaspoon vanilla extract
- 2 large eggs
- 2 cups NESTLE® TOLL HOUSE® Semi-Sweet Chocolate Morsels
- 1 cup chopped nuts

Instructions: Preheat oven to 375°F. Combine flour, baking soda and salt in small bowl. Beat butter, granulated sugar, brown sugar and vanilla extract in large mixer bowl until creamy. Add eggs, one at a time, beating well after each addition. Gradually beat in flour mixture. Stir in morsels and nuts. Drop by rounded tablespoon onto ungreased baking sheets. Bake for 9 to 11 minutes or until golden brown. Cool on baking sheets for 2 minutes; remove to wire racks to cool completely.

Cookie A- Gradient Boosting

Ingredients

- 1.5 c butter
- 2 egg
- .25 c sugar
- 1 tsp vanilla
- 2 c flaked coconut
- 12 tsp hot water
- 0.5 mashed avocado
- 0.75 c matzo cake meal
- 0.25 plain yogurt
- 1 tsp salt
- 16 oz semisweet choc chips

Instructions: startseq whisk together the flour and butter add confectioners sugar and vanilla extract with an an electric add whisk flour and stir until dough is distributed and enough to least least least least least thirty minutes or until not break up the not not break up the not half add chocolate and not not half add chocolate and not not half add chocolate and not not half add chocolate and not not half enough to least least least minutes or until not not not not not not not not not not not not not not not not not not not not not not half add chocolate and not not half add chocolate enough to separate

Cookie B- Extreme Tree

Ingredients

- 1 tsp baking soda
- 1 c brown sugar
- 1 c butter
- 2 egg
- 1.75 c flour
- 1.5 c mint filled morsels
- 0.25 c sugar
- 1 tsp vanilla
- 1 tsp baking powder
- 1 egg yolk
- 1 tsp ground cinnamon
- 0.5 tsp salt
- 0.5 c shortening

Instructions: startseq preheat oven to three hundred and fifty degrees one hundred and seventy five degrees in medium bowl cream together the butter brown sugar and white sugar until smooth beat in the eggs one at time then stir in the vanilla and vanilla combine the flour baking soda and salt stir into the creamed mixture until just blended fold in the chocolate chips drop by rounded spoonfuls onto the prepared cookie sheets bake for eight to ten minutes in the preheated oven allow cookies to cool on baking sheet for five minutes before removing to wire rack to cool completely endseq

Cookie C- Extreme Tree

Ingredients

- 1 tsp baking soda
- .75 c brown sugar
- .5 c butter
- 4 eggs
- 4 c flour
- 4 oz creamy PB
- 2 tsp ground cinnamon
- .5 c mashed avocado
- .66 c milk chocolate chips
- 0.5 tsp salt
- 12 oz semi sweet choc chips

Instructions: startseq preheat oven to three hundred and seventy five degrees one hundred and ninety degrees in medium bowl whisk together the butter brown sugar and brown sugar with an electric mixer in large bowl until smooth add eggs one at medium speed beat in the eggs one at time beating each addition beat in the flour mixture stir in the chocolate chips and walnuts roll balls and place two inches apart on ungreased cookie sheet bake for eight to ten minutes in the preheated oven allow cookies to cool on baking sheet for five minutes before removing to wire rack to cool completely endseq

Cookie D: Deep Learning

Ingredients:

- 1 tsp baking soda
- 1.75 c butter
- 2 eggs
- 1.25 c flour
- .33 c sugar
- .25 tsp vanilla
- .66 c cocoa powder
- 5.28 oz creamy pb
- 2 egg yolk
- 5.28 tbsp espresso powder
- 5 tsp salt
- 8 oz semisweet

Instructions: startseq preheat oven to three hundred and fifty degrees one hundred and seventy five degrees in medium bowl whisk together the butter brown sugar and white sugar until smooth beat in the eggs one at time then stir in the vanilla combine the flour baking soda and salt stir in the chocolate chips and walnuts roll dough into balls and place two inches apart on ungreased baking sheet bake for eight to ten minutes in the preheated oven allow cookies to cool on baking sheet for five minutes before removing to wire rack to cool completely endseq

Cookie E: Deep Learning *modified for baking*

This cookie included no wet ingredients such as butter or eggs to bind the ingredients together and would have just been a pile of dry ingredients, not a cookie. Thus, a stick of butter and 2 eggs were added, plus just enough water for the dough to stick together nicely. Ingredients

- 4 tsp Baking Soda
- 1 c brown sugar
- 4 c flour
- .25 c sugar
- 1 tsp vanilla
- 3 c confectioners sugar
- 16 oz semisweet choc chips
- 5 c walnuts

Instructions: startseq preheat oven to three hundred and seventy five degrees one hundred and ninety degrees in medium bowl whisk together the butter brown sugar and white sugar with an electric mixer in large bowl until smooth add one whisk in the eggs one whisk in separate bowl whisk together the flour mixture and add chocolate and chocolate and not not not combine place balls place one inch balls place one inch balls place two inches balls place two inches balls place one inch balls place one inch balls place one inch balls place balls place one inch balls place one inch balls place balls ball ball ball bake in the preheated oven until set about ten minutes endseq

Procedure for Cookie Tasting[1]

By Dr. Pablo Rivas (Advisor)
For Mackenzie O'Brien (Honors Student)

Please follow this steps for executing the cookie tasting experiment:

1. Explain the participant the risks associated with this experiment and make available to the participant a printed copy of the document entitled "Waiver and Release to Medical Attention and Grant of Rights" for further reading and be ready to answer any questions before proceeding.
2. Provide each participant two cookies: the experimental cookie and the control cookie. Samples can be placed in small bag labeled, for example, A and B.
3. Every bag will contain labels with QR codes and links to an online survey for each cookie, and the table will have a napkins and a cups of water available for each participant.
4. Participants should not talk to one another during the tasting event and should not be able to see how other participants are scoring each sample.
5. Ask each participant to test sample "A" by first recording their score for appearance, then aroma, then taste, and finally texture. Note that texture pertains to how the food feels in your mouth. For example: crunchy, chewy, juicy, soggy, creamy, and so on.
6. After testing the sample, have participants drink water to cleanse their palate.
7. Repeat steps five and six for sample "B."

[1] Based on this:

Waiver And Release Consent to Medical Attention And Grant of Rights[1]

By participating in the cookie tasting and in consideration for my being allowed to participate, the receipt and sufficiency of which is hereby acknowledged, I agree to be bound by each of the following provisions of this waiver, release of liability, indemnification, consent to medical attention and grant of rights (“Waiver”) :

1. Voluntary Participation. I understand and confirm that my participation in the cookie tasting experiment is voluntary. I am in good health and suffer from no food allergy that would make me especially susceptible to injury or disability while participating in the cookie tasting experiment.

2. Comprehension of Risk. I fully comprehend and accept all of the risks associated with my participation in the cookie tasting experiment including, without limitation, injury or death resulting from food sickness, allergic reactions, and choking. I understand that the cookie tasting experiment takes place in public venues under conditions largely beyond our control.

3. Assumption of Risk. Participant fully comprehends and accepts all of the risks associated with his/her participation in the cookie tasting experiment including, without limitation, food sickness, and death.

4. Release of Liability; Limitation of Damages. PARTICIPANT’S PARTICIPATION IN THE COOKIE TASTING EXPERIMENT IS AT PARTICIPANT’S OWN SOLE RISK. PARTICIPANT, ON BEHALF OF HIS/HERSELF AND/OR ANY PERSON OR ENTITY ACTING THROUGH OR ON BEHALF OF PARTICIPANT, HEREBY FOREVER AND UNCONDITIONALLY RELEASES MARIST COLLEGE, AND ANY STUDENTS, PROFESSORS, AND ANY EMPLOYEES, FROM ANY AND ALL CLAIMS, ACTIONS, DAMAGES, LIABILITIES, LOSSES, COSTS AND EXPENSES IN ANY WAY ARISING OUT OF, OR RESULTING FROM, PARTICIPANT’S PARTICIPATION IN THE COOKIE TASTING EXPERIMENT, INCLUDING, WITHOUT LIMITATION, ANY AND ALL CLAIMS,

---

[1] Based on this work: https://tasteittours.com/waiver-release/

ACTIONS, AND LIABILITIES FOR DEATH, INJURY, LOSS OR DAMAGE TO PARTICIPANT, TO ANYONE ELSE, OR TO ANY PROPERTY, REGARDLESS OF WHETHER OR NOT SUCH INJURY, LOSS OR DAMAGE WAS CAUSED BY THE NEGLIGENCE OR WILLFUL CONDUCT OF THE RESEARCHER OR ANY OF THE RELEASED PARTIES. PARTICIPANT, ON BEHALF OF HIS/HERSELF AND/OR ANY PERSON OR ENTITY ACTING THROUGH OR ON BEHALF OF PARTICIPANT, FURTHER AGREES TO DEFEND AND INDEMNIFY THE RELEASED PARTIES, AND TO HOLD THE RELEASED PARTIES HARMLESS, FROM ANY AND ALL LIABILITIES, CLAIMS, ACTIONS, DAMAGES, EXPENSES (INCLUDING, WITHOUT LIMITATION, ATTORNEY'S FEES) AND LOSSES OF ANY KIND OR NATURE WHATSOEVER IN ANY WAY ARISING OUT OF, OR RESULTING FROM, PARTICIPANT'S PARTICIPATION IN THE COOKIE TASTING EXPERIMENT.

5. Consent to Medical Treatment. I authorize Marist College to provide to me, through medical personnel of its choice, customary medical assistance, transportation, and emergency medical services. This consent does not impose a duty upon Marist College to provide such assistance, transportation, or services.

6. Severability. If any provision of this Waiver is for any reason declared to be invalid or unenforceable, the validity and enforceability of the remaining provisions will not be affected. The invalid or unenforceable provision will be modified to the extent necessary to render it valid and enforceable, and if no modification may render it valid and enforceable, this Waiver will be construed as if not containing such provision and the rights and obligations of the parties will be construed and enforced accordingly.

7. Governing Law and Venue. This Waiver shall be governed in all respects by the laws of New York without regard to conflict of law principles. Venue shall be in Poughkeepsie, New York.

None of the provisions of this Waiver and Release of Claims can be waived or modified except expressly in writing signed by Participant and the party against whom the waiver or modification is sought to be enforced. Failure of any of the Released Parties to enforce any of their rights hereunder at any time shall not act as a waiver to enforce their rights under this Waiver and Release for same or similar acts at any subsequent time.

Batch A Cookie 2

* Required

Understand the risks before you continue

If you have not done so, please make sure you have read the risks and understood your rights. You should have been provided with a document entitled "Waiver And Release Consent To Medical Attention And Grant Of Rights" that you need to read. After reading it, please answer the following questions.

1. **Do you understand the risks associated with this research and consent to participate in this research voluntarily?** *

   Mark only one oval.

   - Yes, I understand the risks and I consent to voluntarily participate
   - No, I do not wish to participate at this time

About the Survey

This survey is intended for academic research. As such, your participation is appreciated, but not mandatory. Your responses will be added to others and your identity and participation will remain confidential. This survey has a total of 10 questions and it should take you about 2 minutes to complete. This survey includes demographic questions.
Only adults can participate in this survey. If you are less than 18 years old, please do not participate.

2. **As consenting adult do you agree to respond to this survey in all honesty and truthfulness to the best of your ability?** *

   Mark only one oval.

   - Yes
   - No

Skip to question 3.

Rate the Cookie

In the next three categories, please rate the cookie using the following criteria:
1 = Not fit for consumption
2 = Poor
3 = Neutral
4 = Good
5 = Excellent

3. **Appearance** *

   Mark only one oval.

	1	2	3	4	5	
Not fit for consumption						Excellent

4. **Aroma** *

   Mark only one oval.

	1	2	3	4	5	
Not fit for consumption						Excellent

5. **Taste** *
Mark only one oval.

	1	2	3	4	5	
Not fit for consumption	◯	◯	◯	◯	◯	Excellent

6. **Texture** *
Check all that apply.

☐ crunchy
☐ chewy
☐ gooey
☐ juicy
☐ soggy
☐ creamy
☐ Other: ____________________

7. **On a scale of 1-10, how would you rate your overall satisfaction with this cookie?** *
Mark only one oval.

	1	2	3	4	5	6	7	8	9	10	
Hated It	◯	◯	◯	◯	◯	◯	◯	◯	◯	◯	Loved It

8. **It is important for us to be able to reach you if we need to. Please give us your valid marist email address:** *

____

9. **If you want, we could let you know the next time a tasting event will take place. Would you like that?** *
Mark only one oval.

◯ Yes, email me the details
◯ No, I'm fine

Demographics

10. **What is you gender?** *
Mark only one oval.

◯ Female
◯ Male
◯ Other: ____________________

11. **How old are you?** *

Mark only one oval.

- ( ) 18-24
- ( ) 25-34
- ( ) 35-44
- ( ) 45-54
- ( ) 55-64
- ( ) 65-74
- ( ) 75+

Powered by
Google Forms

Optimizing Cookie Recipes for Ratings Using Machine Learning and Deep Vector-to-Sequence Recurrent Neural Models

Mackenzie O'Brien and Dr. Pablo Rivas

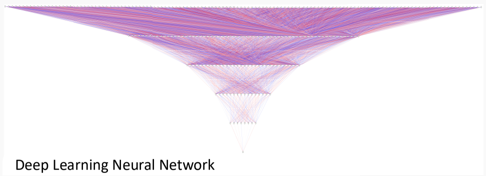
Our Idea

This project employed **machine learning** concepts and processes in an attempt to teach a computer an algorithm that outputs a **cookie recipe** that is comparable to human-made recipes.

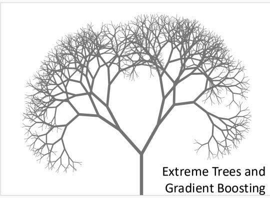
The process involved showing the algorithms a data set of cookie recipes that included details such as ingredients, ratings, and calories. A **deep LSTM** algorithm was used to construct instructions for a recipe once ingredients were selected based off of instructions from the dataset of recipes.

The recipes were gathered by scraping the website allrecipes.com for chocolate chip cookie recipes that were compiled in order to "teach" the computer **to recognize and create** tasty cookies by looking at tactors such as reviews, ingredients, and calories.

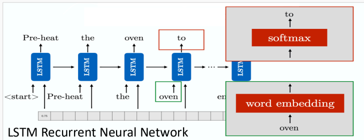
Machine Learning Architectures



Deep Learning Neural Network



Extreme Trees and Gradient Boosting



LSTM Recurrent Neural Network

References

Clifford, Catherine. "Alphabet Billionaire Eric Schmidt: Google Used A.I. to Find the Perfect Chocolate Chip Cookie Recipe." CNBC, CNBC, 29 Jan. 2018.
Kochanski, Greg et al. "Bayesian Optimization for a Better Dessert." (2017).
Naik, JitendraB.. "Cuisine Classification and Recipe Generation." (2015).

Methodology

Gathering Recipes for Data Input

250 chocolate chip cookie recipes were taken from allrecipes.com. This data was then cleaned and used to create ingredients and instructions.

Creating Training Models For Ingredients Vectors

Deep Learning Neural Network (DNN)

This a nonlinear artificial neural network with multiple layers between the input and output layers. Each mathematical manipulation of the data is considered a layer, and complex DNN have many layers.

Extremely Randomized Trees

The algorithm is a variant of the Random Forest regression that picks a decision boundary at random by using the entire sample at each step of the extreme tree.

Gradient Boosting

The gradient boosting algorithm applies the concept of modifying weak learners to become better at identifying good outcomes more efficiently.

ML Error Rating Prediction via Mean Square Error

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^*)^2$$

Vector-to-Sequence Models to Get Instructions

Long Short-Term Memory Recurrent Neural Network

We used each generated recipe as an attention mechanism, which when combined with LSTM, trains the model to output a sequence of words that it believes are good recipe.

Minimizing Cross Entropy Loss

$$\frac{1}{N}\sum_{n=1}^{N}\sum_{c\in C} d_{cn}^* \ln d_{cn} + (1 - d_{cn}^*)\ln(1 - d_{cn})$$

Testing the Cookies

After evaluating the models, we chose 5 recipes to test. We did taste tests of each machine learning cookie against the same control cookie, serving 40 of each cookie on campus per test. The typical rate of response was around 30 for each cookie.

Quality Metrics of Cookie Recipes

Simplicity

$||y - y^*|| + ||\mathbf{w}||$

Uniqueness

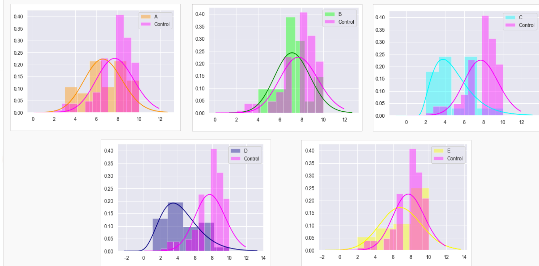
$\min_{i=1,\ldots,N} ||\boldsymbol{x}_i - x^*||$

Survey Questions

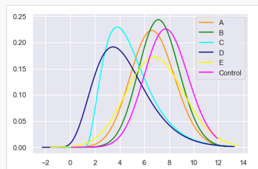
1. Appearance on a scale of 1 (unfit for consumption) to 5 (Excellent)
2. Aroma on a scale of 1 (unfit for consumption) to 5 (Excellent)
3. Taste on a scale of 1 (unfit for consumption) to 5 (Excellent)
4. Texture: crunchy, chewy, gooey, juicy, soggy, creamy, other
5. Overall Satisfaction on a scale of 1 (Hated It) to 10 (Loved It)

Graphs

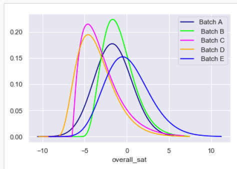
Individual Batches vs Control



All Batches vs Control



Δ: Machine Learning – Control



Cookies

Results

Ingredients

Batch B had the best results, but none of the recipes beat the control cookie. Overall, the recipes generated had good responses during experiment tastings. **Recipe E had to be modified**, as it included no wet ingredients, and thus would not correctly form dough. The **best algorithm** out of those used was **Extremely Randomized Trees.**

Vector-to-Sequence Instructions

Recipe instructions were created, **however some were not viable for use**, and none generated a sequence that included every ingredient in the recipe fed into it. This may be because our dataset was not robust enough, but that can possibly be remedied through using more recipes during training the algorithm in further testing.

Improvements include ensuring that the instructions contain all ingredients in the vector that contain non-zero values and eliminating repeating loops that the algorithm gets stuck on.