Alternate Food Recommendation

Why?

Recommendations are essential because transitioning to meat alternatives can inadvertently reduce vital nutrients—such as protein, iron, and vitamin B12—while diverse, nutrient-dense diets remain critical for health across all life stages.

Gym Freaks

Gym Freaks Regularly seek for change in their diet which can provide them with same amount of nutrition but don’t know what to consume

Dieticians

Dieticians across the world face the problem of suggesting food items that can provide you with appropriate amount of micro as well as macro-nutrients

Farsi, D. N., Uthumange, D., Munoz Munoz, J., & Commane, D. M. (2022). The nutritional impact of replacing dietary meat with meat alternatives in the UK: a modelling analysis using nationally representative data. The British journal of nutrition, 127(11), 1731–1741. https://doi.org/10.1017/S0007114521002750

Problem Statement

As of now there exists no machine learning model which accounts for micro-nutrients while giving the alternative to the food that is providing you with the same amount of nutrition

Micronutrients are very small in quantities so its difficult to account them when comparing macronutrients as well.

C. G. Awuchi, V. S. Igwe, I. O. Amagwula, and C. K. Echeta, "Health Benefits of Micronutrients (Vitamins and Minerals) and their Associated Deficiency Diseases: A Systematic Review," International Journal of Food Sciences, vol. 3, no. 1, pp. 1–32, 2020. [Online]. Available: Kim, S., Fenech, M.F. & Kim, PJ. Nutritionally recommended food for semi- to strict vegetarian diets based on large-scale nutrient composition data. Sci Rep 8, 4344 (2018). https://doi.org/10.1038/s41598-018-22691-1

Why Micronutrients are necessary

Micronutrients deficiencies can impact almost ALL aspects of human health and wellbeing throughout the life course [20]. Perhaps most alarming are the potential impacts on health from even subtle or subclinical deficiencies, which are often also asymptomatic and therefore hidden

Health impacts from these deficiencies are many and include:

- Weakened Immune Response
- Diminished Cognitive Function and Brain Health
- Increased Risks of Non-Communicable Diseases (NCDs)
- Widespread Fatigue

Chen, Y., Michalak, M., & Agellon, L. B. (2018). Importance of Nutrients and Nutrient Metabolism on Human Health. The Yale journal of biology and medicine, 91(2), 95–103.

Solution

Clear and Concise

Micronutrients play a crucial role in maintaining our body's vital functions. They support essential processes like metabolism, and their deficiency can disrupt overall health and well-being.

Accounting micro-nutrients

While training our machine learning model we will account for both macro as well as micro nutrients

Shenkin A. (2006). Micronutrients in health and disease. Postgraduate medical journal, 82(971), 559–567. https://doi.org/10.1136/pgmj.2006.047670

DataSet

The Dataset we chose belongs to USDA (United States Department of Agriculture). The Dataset had more than 500,000 different food items. But we picked a small number from it and made our regression model on 7,000 different foods that are often consumed.

Advantage of this dataset

- accounts for both macro and micro nutrients
- Reliable and Verified Data set provided by the US Government itself

Data pre proccessing

Data Integration

- The provided data was very unstructured for our models consumption.
- We selected Major Macro nutrients and Micro nutrients which affected the food uniqueness
- Macronutrients: Protein, Carbohydrates, Sugars, Fat, Fiber, Energy (KCAL, kJ), and Water.
- Vitamins: Vitamin A (IU, RAE), B-12, B-6, C, D, E, K, Thiamin, Riboflavin
- Minerals: Calcium, Iron, Magnesium, Phosphorus, Potassium, Sodium, Zinc, Selenium, Copper, Manganese.

Data pre proccessing

- We first looked at pictures (like the histograms) of our nutrient data to see how the numbers for each nutrient were spread out.
- We saw that most of the nutrient data was "skewed to the right." This means the numbers were mostly low, but there was a long tail of some high numbers.
- To make the skewed data more balanced, we used a special math step called log transformation. This helped to squeeze the high numbers closer together.
- After fixing the skewness for some data, and for all the data, we adjusted the numbers for each nutrient so they are all on a similar scale. We used two ways to do this: Normalization and Standardization

dataset (final form)

- Protein (G)
- Carbohydrate, by difference (G)
- Sugars, Total (G)
- Fiber, total dietary (G)
- Total lipid (fat) (G)
- Energy (KCAL)
- Energy (kJ)
- Water (G)
- Riboflavin (MG)
- Calcium, Ca (MG)
- Iron, Fe (MG)
- Magnesium, Mg (MG)
- Phosphorus, P (MG)
- Potassium, K (MG)
- Sodium, Na (MG)
- Zinc, Zn (MG)
- Selenium, Se (UG)
- Copper, Cu (MG)
- Manganese, Mn (MG)

- Vitamin A, IU (IU)
- Vitamin A, RAE (UG)
- Vitamin B-12 (UG)
- Vitamin B-12, added (UG)
- Vitamin B-6 (MG)
- Vitamin C, total ascorbic acid (MG)
- Vitamin D (D2 + D3) (UG)
- Vitamin D (D2 + D3), International Units (IU)
- Vitamin E (alpha-tocopherol) (MG)
- Vitamin E, added (MG)
- Vitamin K (phylloquinone) (UG)
- Thiamin (MG)

Litrature Review

- We Studied many papers and studied how people usually do it. In one of many recent researches. They developed a system which could capture the image and tell the nutritional composition of the food and based on that it can recommend similar food item.

- For the recommendation part they used regression model

- To validate the regression model they used MAE, MSE and R(square) scores.

- The biggest gap we found that they gave there recommendation using only macro nutrients

Zhao, Y., Zhu, P., Jiang, Y., & Xia, K. (2024). Visual nutrition analysis: leveraging segmentation and regression for food nutrient estimation. Frontiers in Nutrition, 11. https://doi.org/10.3389/fnut.2024.1469878

Another Research

- We Studied another research paper in which the researchers are using autoencoders and also accounting for user data while recommending them the food

- For the recommendation part they used regression model

- To validate the regression model they used MAE, MSE and R(square) scores.

- Once again they only accounted for the macro nutrients not touched upon the usage of micro nutrients

Ahn, D. (2024). Accurate and Reliable Food Nutrition Estimation Based on Uncertainty-Driven Deep Learning Model. Applied Sciences, 14(18), 8575. https://doi.org/10.3390/app14188575

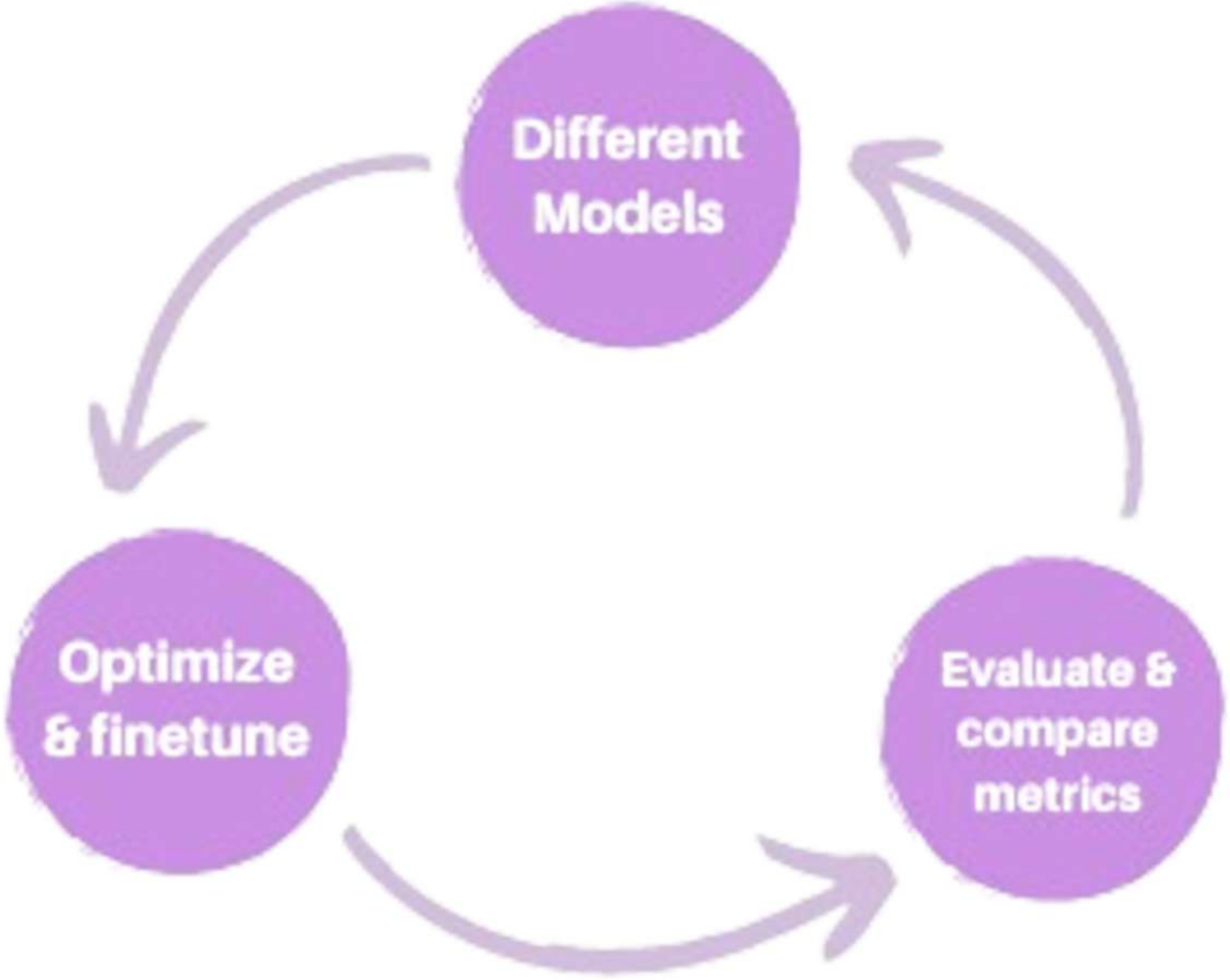
Gap Orignates

- Like these many researchers haven’t used micro nutrients in there research while developing a robust system.
- This is most likely because micro-nutrients are difficult to account for due to there less amount of presence in the food

- With this we found the right dataset and integrated the micro-nutrients in the recommendation systems as well.

Proposed Methodologies

WorkFlow


Different
Models
Evaluate &
compare
metrics
Optimize
& finetune

We tried Many Models

How well they worked?

DB Scan

Silhouette Score: -0.0640

This is poor, indicating that the clustering structure may not be meaningful or structured

.Davies-Bouldin Score: 0.8369

This is better than KMeans in terms of compactness and separation But still not distinguishable clusters are formed

K- Means

Silhouette Score: 0.2242

This is a moderate score. Silhouette ranges from -1 to 1 (Values near 1 indicate well-separated clusters)

.Davies-Bouldin Score: 1.9031

A score of ~1.9 indicates moderate to poor clustering, suggesting that clusters are not very compact and are somewhat close to each other.

Infrences

- As we were playing with DB Scan and K-Means clustering algorithm we were not able to form clusters that were relevent to us so we had to look for other ML methodologies

- We decided to use Autoencoders because they are good for food recommendation based solely on nutrient data, as they can extract compact, non-linear embeddings from high-dimensional, skewed numerical features. This allows for discovering hidden nutritional similarities between foods, enabling meaningful recommendations

Xie, J., Girshick, R., & Farhadi, A. (2015). Unsupervised deep embedding for clustering analysis. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1511.06335

Papastratis, I., Konstantinidis, D., Daras, P. et al. AI nutrition recommendation using a deep generative model and ChatGPT. Sci Rep 14, 14620 (2024). https://doi.org/10.1038/s41598-024-

Auto Encoders

- MSE: 0.0131 - Significantly lower than the denoising autoencoder, indicating a much more precise reconstruction
- MAE: 0.0495 - Lower than the denoising autoencoder, suggesting smaller average absolute deviations from the actual values.
- R^2 Score: 0.9853

Why These are imortant

MSE (Mean Squared Error):
Captures how precisely the autoencoder reconstructs original nutrient profiles.

MAE (Mean Absolute Error):
Captures how precisely the autoencoder reconstructs original nutrient profiles.

R^2 Score (Coefficient of Determination)
Indicates how well the model captures the underlying structure of nutrient data.

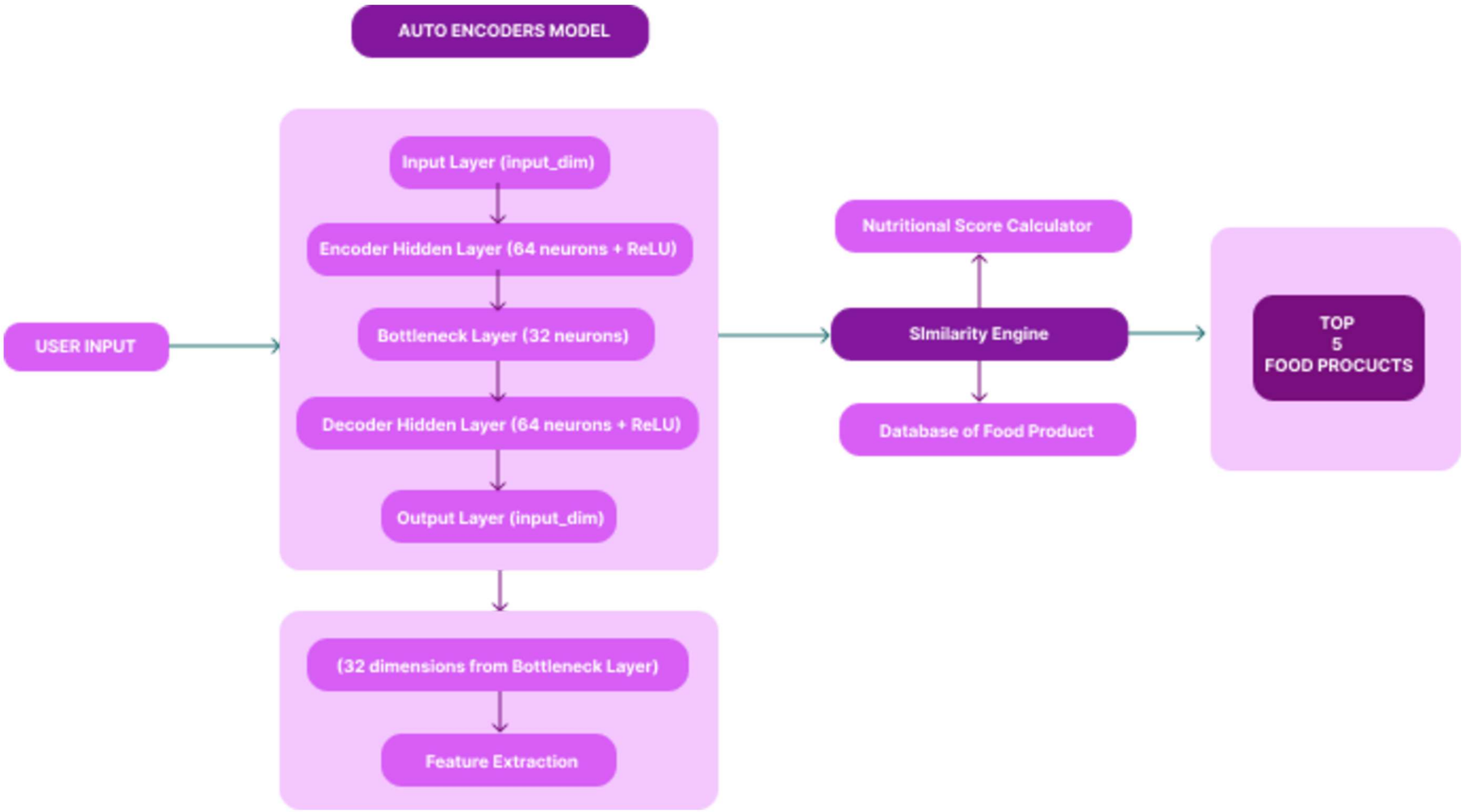
De-noising Auto Encoders

- MSE: 0.0131 - Significantly lower than the denoising autoencoder, indicating a much more precise reconstruction
- MAE: 0.0495 - Lower than the denoising autoencoder, suggesting smaller average absolute deviations from the actual values.
- R^2 Score: 0.9853

MSE (Mean Squared Error):
Captures how precisely the autoencoder reconstructs original nutrient profiles.

MAE (Mean Absolute Error):
Captures how precisely the autoencoder reconstructs original nutrient profiles.

R^2 Score (Coefficient of Determination)
Indicates how well the model captures the underlying structure of nutrient data.


AUTO ENCODERS MODEL
USER INPUT
Input Layer (input_dim)
Encoder Hidden Layer (64 neurons + ReLU)
Bottleneck Layer (32 neurons)
Decoder Hidden Layer (64 neurons + ReLU)
Output Layer (input_dim)
(32 dimensions from Bottleneck Layer)
Feature Extraction
Nutritional Score Calculator
SImilarity Engine
Database of Food Product
TOP
5
FOOD PROCUCTS

Why We Chose Autoencoders Over Denoising Autoencoders

Lower reconstruction error
Autoencoders had lower MSE and RMSE, meaning more accurate nutrient predictions.

Higher R^2 score
Showed that autoencoders captured more variance in the nutrient data.

We explored denoising autoencoders to test if adding noise could help the model learn more robust nutrient patterns and generalize better, especially given the skewed, non-linear nature of our data. It was a way to check if resilience to slight input variation could improve recommendation quality.

- **Structural difference**: Denoising autoencoders add noise to the input before encoding; standard autoencoders use clean input.
- **Functional difference**: Standard autoencoders aim to reconstruct exact input; denoising autoencoders learn to recover clean data from noisy input, improving robustness.

Papastratis, I., Konstantinidis, D., Daras, P. et al. AI nutrition recommendation using a deep generative model and ChatGPT. Sci Rep 14, 14620 (2024). https://doi.org/10.1038/s41598-024-65438-x

Challenges Faced:

- Our biggest challenge was to find the right Dataset. First we explored and found a dataset on kaggle, after using it and training a model we were disappointed because the data usability was very low. Many of the fields were empty even though the food had that particular mineral. Then we stumbled across a very good dataset from a reputative organization USDA.

- Next challenge we faced was to look for right ML methodologies. First we thought with the help of clustering we could easily achieve it. But, clustering algorithm like K-means and DB-Scan failed and gave us bad results. Then we started searching what other people are executing in the similar field, we happen to stumble across auto-encoders and it gave us satisfactory results.

references

Lu, P. M., & Zhang, Z. (2025). The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering. Journal of Computational Biology and Medicine, 5(1).
accounts macronutrients and KNN algorithm

Darmon, N., Sondey, J., Azaïs-Braesco, V. et al. The SENS algorithm—a new nutrient profiling system for food labelling in Europe. Eur J Clin Nutr 72, 236–248 (2018). https://doi.org/10.1038/s41430-017-0017-6

Shenkin A. (2006). Micronutrients in health and disease. Postgraduate medical journal, 82(971), 559–567. https://doi.org/10.1136/pgmj.2006.047670

C. G. Awuchi, V. S. Igwe, I. O. Amagwula, and C. K. Echeta, "Health Benefits of Micronutrients (Vitamins and Minerals) and their Associated Deficiency Diseases: A Systematic Review," International Journal of Food Sciences, vol. 3, no. 1, pp. 1–32, 2020. [Online]. Available: Kim, S., Fenech, M.F. & Kim, PJ. Nutritionally recommended food for semi- to strict vegetarian diets based on large-scale nutrient composition data. Sci Rep 8, 4344 (2018). https://doi.org/10.1038/s41598-018-22691-1

Farsi, D. N., Uthumange, D., Munoz Munoz, J., & Commane, D. M. (2022). The nutritional impact of replacing dietary meat with meat alternatives in the UK: a modelling analysis using nationally representative data. The British journal of nutrition, 127(11), 1731–1741. https://doi.org/10.1017/S0007114521002750

Papastratis, I., Konstantinidis, D., Daras, P. et al. AI nutrition recommendation using a deep generative model and ChatGPT. Sci Rep 14, 14620 (2024). https://doi.org/10.1038/s41598-024-65438-x

Thank You