Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.034324
Article

Weight Prediction Using the Hybrid Stacked-LSTM Food Selection Model

Ahmed M. Elshewey[1], Mahmoud Y. Shams[2,*], Zahraa Tarek[3], Mohamed Megahed[4], El-Sayed M. El-kenawy[5] and Mohamed A. El-dosuky[3,6]

[1]Faculty of Computers and Information, Computer Science Department, Suez University, Suez, Egypt
[2]Faculty of Artificial Intelligence, Kafrelsheikh University, Kafrelsheikh, 33511, Egypt
[3]Faculty of Computers and Information, Computer Science Department, Mansoura University, 35561, Egypt
[4]Faculty of Physical Education, Track and Field Competitions Department, Arish University, Egypt
[5]Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt
[6]Department of Computer Science, Arab East Colleges, Riyadh, 13544, Saudi Arabia
*Corresponding Author: Mahmoud Y. Shams. Email: mahmoud.yasin@ai.kfs.edu.eg
Received: 13 July 2022; Accepted: 18 August 2022

Abstract: Food choice motives (i.e., mood, health, natural content, convenience, sensory appeal, price, familiarities, ethical concerns, and weight control) have an important role in transforming the current food system to ensure the healthiness of people and the sustainability of the world. Researchers from several domains have presented several models addressing issues influencing food choice over the years. However, a multidisciplinary approach is required to better understand how various aspects interact with one another during the decision-making procedure. In this paper, four Deep Learning (DL) models and one Machine Learning (ML) model are utilized to predict the weight in pounds based on food choices. The Long Short-Term Memory (LSTM) model, stacked-LSTM model, Conventional Neural Network (CNN) model, and CNN-LSTM model are the used deep learning models. While the applied ML model is the K-Nearest Neighbor (KNN) regressor. The efficiency of the proposed model was determined based on the error rate obtained from the experimental results. The findings indicated that Mean Absolute Error (MAE) is 0.0087, the Mean Square Error (MSE) is 0.00011, the Median Absolute Error (MedAE) is 0.006, the Root Mean Square Error (RMSE) is 0.011, and the Mean Absolute Percentage Error (MAPE) is 21. Therefore, the results demonstrated that the stacked LSTM achieved improved results compared with the LSTM, CNN, CNN-LSTM, and KNN regressor.

Keywords: Weight prediction; machine learning; deep learning; LSTM; CNN; KNN

1 Introduction

Food choices have a significant impact on the long-term viability of modern diets; therefore, knowing consumer dietary patterns is critical in accomplishing sustainability objectives. Understanding food choices is difficult due to the wide range of elements that influence them, including socio-demographic characteristics, attitudes, beliefs, conventions, consumption environments, and cultural settings. Beyond the aforementioned elements, insights into food decision reasons are useful in comprehending food choices. Food choice motives are the reasons or motives that customers have for selecting or eating food [1]. In the context of long-term food exhaustion, the significance of insights into consumers' food choice reasons has been demonstrated. Some reasons, e.g., concern for the environment, might be regarded as supportive of sustainable food choices, while others may be seen as roadblocks to making sustainable dietary choices (as, when sustainable food is less delicious or less adequate). Furthermore, various underlying motives can be used to justify various types of sustainable eating choices. Because of the possibility for disputes and trade-offs among varied dietary preferences, it is crucial to investigate sustainability reasons across a wide range of conceivable food-choice factor [2]. Obesity and overweight are significant risk factors for some chronic diseases, such as cancer, cardiovascular, and diabetes disease. Overweight is a public health issue with limited pharmaceutical treatments. It is unknown if individuals with obesity may lose weight with semaglutide at a dosage of 2.4 mg once a week as well as lifestyle adjustments [3]. Variation in obesity incidence among nations may be described by the 'caloric ecosystem dynamics': elements affecting food production, consumption, distribution, culture of food, and habits on a wide scale [4]. Finding the link between obesity and food sales is a huge task that has recently been called into doubt. Traditional regression procedures restrict the study to a narrow range of factors and apply independence and linearity assumptions. When it comes to predicting the impact of variations in national diets made up of closely related food categories, these assumptions are virtually likely broken. Instead, we will look at a machine learning (ML) technique, which, unlike first-principles approaches, does not require the analyst to describe a functional form of the model [5]. Machine learning systems employ training samples to uncover patterns and, eventually, to accurately predict future trends. Traditional machine learning methods have often been insufficient with the development of machine learning methodologies, as well as the scale and complexity of data processing. Accordingly, deep learning algorithms have been created, which have more complex structures and greater power in data analysis. Deep learning, also known as a deep neural network (DNN), is a learning method that involves recursive processes in several layers to develop a deep structure and produce a model. Although the interpretation of deep learning is inferior to the interpretation of classical machine-learning approaches, its learning performance is greater [6]. Over time, scientists from many fields have generated issues affecting food choice in conceptual models, which have been identified as a critical lever for improving human and planetary health. However, an interdisciplinary strategy is necessary to understand how different aspects interact with each other in the decision-making process [7]. In this paper, we find that integrating ML and DL techniques in global health can improve predictions and explore rich combinations among the available data, thus improving our understanding of health challenges and aiding the development of new policies. Especially in the areas of nutrition and obesity, the contribution of this paper is the use of the ML method, that is, the K-Nearest Neighbor (KNN)), and four deep learning models are the Long Short-Term memory model (LSTM), the stacked-LSTM model and the Convolutional Neural Network model (CNN) and the CNN-LSTM model to predict pound weight from food selection data. The remainder of this paper is laid out as follows. The literature review is covered in Section 2. The proposed methodology is described in Section 3. Findings and discussions are provided in Section 4. Section 5 presents the conclusions of this study.

2 Related Work

Bi et al. [8] presented a deep learning approach that employed an Auto-Encoder to acquire product attributes from expert-scored sensorial traits, and the sensory data gained was retreated on customer needs using SVM analysis. To verify the total learning process, feature grouping, hedonic contour mapping, and model performance evaluation were used. The findings demonstrated that the DL approach can provide an adequate level of precision, and the hedonic mapping displayed might be of considerable use to manufacturers' product design or adjustments. Pellegrini et al. [9] looked at weight changes and eating patterns in a group of obese outpatients following a month of compulsory confinement during the COVID-19 epidemic in Northern Italy. By completing the questionnaire, a regression analysis with several variables was done to investigate the connections between BMI/weight modifications and the studied factors. They concluded that less exercise, solitude/boredom, depression/anxiety, increased eating, snacking, deleterious meals, grains, and dessert were all associated with considerably greater weight increasing. Hintze et al. [10] examined the effects of changes in Reinforcing Relative Value (RRV) of impulsivity and food could anticipate weight and structure variations in women who are overweight or obese participating in either quick or gradual weight reduction plans. The findings showed that varying rates of weight reduction had no effect on RRV snacking or impulsivity. Nevertheless, variations in RRV snack anticipated weight or fat mass (FM) decreased, indicating that dietary therapies that either reduce or encourage RRV snack decreases produced a greater reduction in obese. Gere et al. [11] provided a classification approaches for food preference prediction. Thirteen classification models were developed and evaluated after variable filtering. The models were compared using the approach of the summing of ranking variances based on performance metrics. The approach categorized the algorithms by matching the ranks of their evaluation matrix to a specified gold criterion. In each case, decision tree-based approaches outscored all others, regardless of the selection tasks or food product groupings. Between the classifiers, the cost-sensitive decision tree algorithm and Quinlan's C4.5 achieved higher accuracy than other approaches. Depa et al. [12] studied food-choice motivations in relation to both features of Orthorexia. They investigated the association between Orthorexia, age, food-choice reasons, sexuality, and BMI. The motivations that predicted dietary preferences in healthy orthorexia and orthorexia-nervosa were very dissimilar. Therefore, orthorexia-nervosa, weight control with sensory resumption and mood regulation were the main motivations that exhibited substantial connections. The major motivation for healthy orthorexia was health context, with sensory resumption and price also exhibiting considerable relationships between the lends credence to the concept that Orthorexia nervosa is linked to maladaptive eating habits that are driven by weight control rather than health issues. Nogales et al. [13] presented a study for predicting food and feed risk concerns. Deep learning combined with entity embedding proved to be the optimum collection, with an accuracy reached to 86.81%, 82.31%, and 88.94% throughout each of the simplified Rapid Alert System for Food and Feed (RASFF) process' three steps in which the experiments were performed. Nevertheless, the Random Forest (RF) systems using a single encoding provided only inferior outcomes, suggesting that the coding had a greater impact on the quality of the findings than the forecasting approach. Table 1 illustrates some of the machine learning and deep learning techniques for different aspects of prediction based on food choice. In the next section, we will discuss the recommended ML/DL techniques for the prediction of weight in pounds based on food choice in the proposed methodology. Moreover, in [14–20], ML approaches are presented to predict suitable food selection with respect to individual choice, food environment, and preference, and obesity, regulation during food choice, food prices, supply chain risk prediction, and crop yield prediction. Shams et al. [21] present a healthy nutrition analysis, as they recommended a suitable diet during the COVID-19 pandemic.

Table 1: ML/DL techniques for prediction of food choice

Author	Techniques	Results	Purpose
Dunstan et al. [5]	Random Forest (RF)-Support Vector Machine (SVM)-eXtreme Gradient Boosting (XGB)	RMSE(SVM) = 0.063 RMSE(RF) = 0.057 RMSE(XGB) = 0.058	Predicting obesity from food
Zhu et al. [14]	Machine Learning-based Classification (MLC)	Accuracy = 74.7% & AUC-ROC = 0.85.	Prediction of individual food choice

(Continued)

Table 1 (continued)

Author	Techniques	Results	Purpose
Amin et al. [15]	Least Absolute Shrinkage and Selection Operator (LASSO)-Random Forests (RF)-eXtreme Gradient Boost (XGB)	Accuracy = 72%	Prediction of the modified Retail Food Environment Index (food deserts and swamps)
Nilashi et al. [16]	Latent Dirichlet Allocation (LDA)-Self Organizing Map (SOM)-Classification and Regression Trees (CART)	MAE = 0.3852, RMSE = 0.4691, MAPE = 8.77% and R^2 = 0.9301	Preference prediction
Cosme et al. [17]	Support Vector Machine (SVM)-Logistic Regression (LR)-Ridge Regression (RR)	Accuracy (LR, RR) = 0.82, Accuracy (SVM) = 0.83	Assessing spontaneous regulation during food choice
Hidayat et al. [18]	Multiple Linear Regression (MLR)	MSE (rice) = 2171.04, MAE = 145.79, RMSE = 145.812, MAPE = 0.81%	Predicting food prices
Baryannis et al. [19]	Support Vector Machine (SVM)-Restricted Decision Trees (RDT)-unrestricted Decision Trees (DT)	Accuracy (SVM) = 0.943 Accuracy (DT) = 0.950 Accuracy (RDT) = 0.916	Supply chain risk prediction
Bian et al. [20]	Support Vector Machine Regression (SVR)-Gaussian Process Regression (GPR)-Random Forest Regression (RFR)	R2 (GPR) = 0.88, MAE = 42.57 g/m^2, RMSE = 49.18 g/m^2	Prediction of crop yield

3 Methodology

Finding the link between food choices and weight from basic principles is a huge task that has recently been called into doubt. Traditional regression procedures restrict the study to a narrow range of factors and apply independence and linearity assumptions [22]. When modelling the influence of disparities in national dietary habits composed of strongly associated food groups, these assumptions are virtually likely broken. Alternatively, we will look at a Machine Learning (ML) technique, which, unlike first-principles approaches, does not require the investigator to describe a functional method of the approach [23]. Deep Learning (DL) models display remarkable capabilities in classification/regression tasks when appropriate data representing the relevant problem is available. Deep learning approach has begun to be employed in the area of dietetics due to its high power of automatic feature learning, mostly for food calories computation, vegetable and fruit quality assessment, and classification of food categories, and so on [24]. A food recommendation using heuristics and ontology is presented in [25]. Fig. 1 shows the system of proposed methodology using different ML and DL techniques to predict weight from food choice. Machine learning is a subset of artificial intelligence that entails models capable of collecting relevant data from input and using that information for self-learning to make accurate categorization or prediction. Machine learning has grown in prominence because of its accuracy and dependability. Artificial Neural Network (ANN), Decision Trees (DT), Fuzzy Logic (FL), k-means clustering, Nave Bayes (NB), Random Forest (RF), K-Nearest Neighbor (KNN) and Support Vector Machines (SVM). In this paper, the

K-Nearest Neighbor (KNN) regressor as a machine learning approach is proposed for the prediction of weights in pound based on food choices.

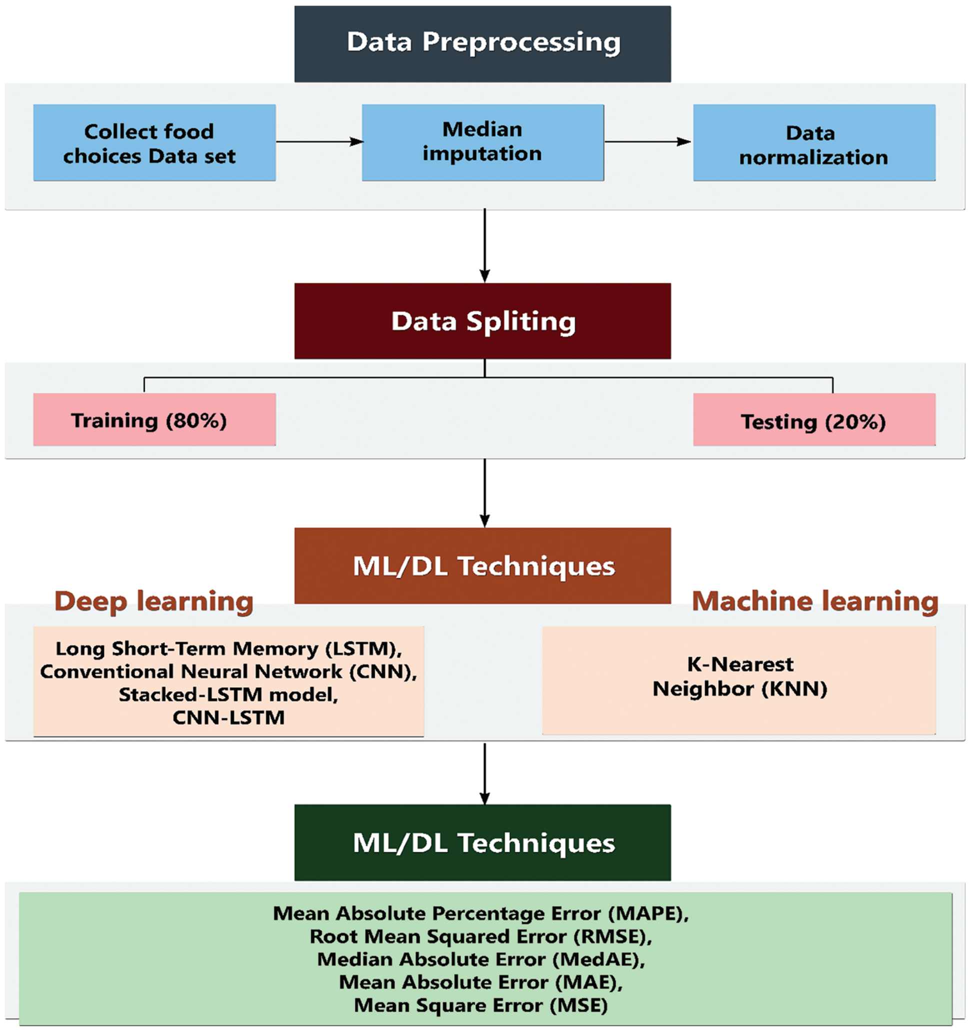

Figure 1: System of the proposed methodology

3.1 K-Nearest Neighbors (KNN)

The distance between various trait values is calculated using KNN to execute classification. The idea is that if the majority of the K similar samples in the trait space (the feature space's nearest neighbors) belong to a particular group, then the sample does as well. K is generally an integer less than or equal to 20. The chosen neighbors in the KNN approach are all samples that have been accurately categorized. Based on the categorized choice of the nearest samples, this technique solely determines the class to which the samples to be categorized belong. To circumvent the matching problem between items, the distance between items is calculated using KNN as a non-similarity index between objects. The distance is usually the Euclidean or Manhattan distance (defined in Eqs. (1) and (2), respectively) [6].

$$\text{Euclidean distance: } d(x,\ y) = \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} \tag{1}$$

$$\text{Manhattan distance: } d(x,\ y) = \sqrt{\sum_{k=1}^{n} |x_k - y_k|} \tag{2}$$

The KNN algorithm is presented in Fig. 2, where it is also shown that the result largely relies on the K value choice. Another machine learning subfield, deep learning, has shown greater performance in the image categorization of various food items and has demonstrated its ability to surpass even humans in certain circumstances [21–25]. In this study, four deep learning models are utilized to predict the weight in pounds based on food choices, namely, the long short-term memory (LSTM) model, conventional neural network (CNN) model, stacked-LSTM model, and CNN-LSTM model.

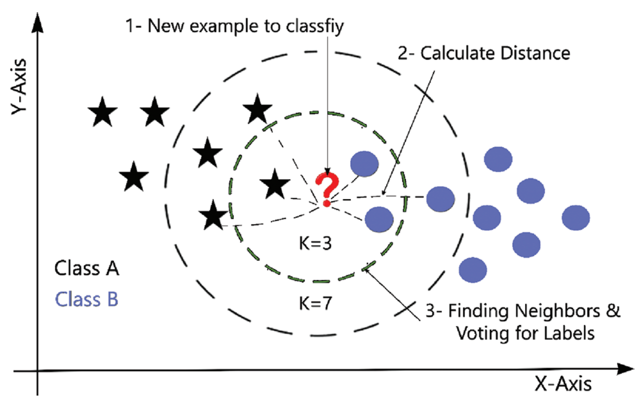

Figure 2: An example of KNN model

3.2 Convolutional Neural Networks (CNN)

Convolutional layers, fully connected layers, pooling layers, nonlinear layers, and the output are the components of CNN. The possibility that precisely captures a particular or group of image groupings is CNN's outcome. CNN image recognition method Fig. 3. Datasets are regarded as necessary for successful training because CNN can learn how to extract relevant features from input images [26]. The input image may be read as a series of matrices, with the output determining what item this image is most probably to be. The dot product of the source images and the weight matrix of the filter is computed by the convolutional layer. The result is utilized as the layer's outcome. The filter will move around the whole image, doing the identical dot product process each time [27,28]. The convolution process may be described as in Eq. (3).

$$s(i,\ j) = (X * W)(i,\ j) + b = \sum_{k=1}^{n\_in} (X_k * W_k\ (i,\ j) + b) \tag{3}$$

where n_in is the no of incoming indices or the tensor's ultimate dimension, X_k indicates the kth source index, W_k is the kth sub-convolution kernel matrix, and s(i, j) is the value of the relevant result matrix element associated with the convolution kernel W. The ReLU function, f(x) = max(0, x), is commonly employed for the output of nonlinear layers. For each negative value in the input image, the ReLU function returns a value of zero, but for each positive value, it returns the same value [6].

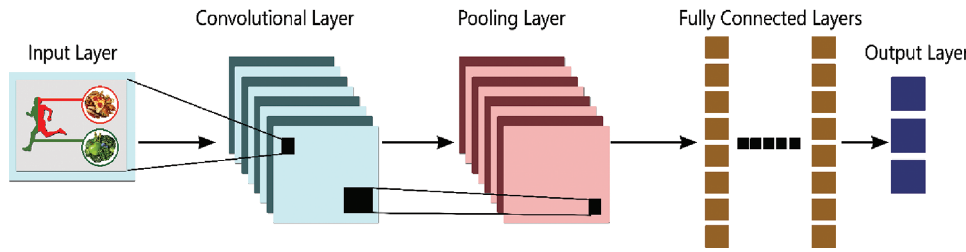

Figure 3: The typical CNN architecture

3.3 Long Short-Term Memory (LSTM)

To natively include time dependency into neural network topologies, the recurrent neural networks have been proposed. The key concept is to extend the capabilities of traditional ANN approaches by dynamically providing a sequential framework. Because of LSTM ability to learn both long-and short-term interconnections of a problem, it has been proven a huge success. It has been also developed to deal with the vanishing gradient issue that most RNN designs have [29]. The primary data processing units of LSTM are referred to as "cells." In normal MLP, these cells may be seen as more advanced neurons. A cell has numerous gates that maintain and govern the flow of information throughout an arbitrary length sequence. This capability allows LSTM to determine whether information is important in both the short and long term. As a result, it is ideal for any form of sequential issue. A LSTM cell is defined as in Eqs. (4)–(9) [30].

$$i_t = \sigma(W_i\,[h_{t-1};\ x_t] + b_i) \tag{4}$$

$$f_t = \sigma\left(W_f\,[h_{t-1};\ x_t] + b_f\right) \tag{5}$$

$$g_t = \tanh(W_s\,[h_{t-1};\ x_t] + b_s) \tag{6}$$

$$o_t = \sigma(W_o\,[h_{t-1};\ x_t] + b_o) \tag{7}$$

$$s_t = f_t \odot s_{t-1} + i_t \odot g_t \tag{8}$$

$$h_t = o_t \odot \tanh(s_t), \tag{9}$$

where x and h represent the input state and the hidden state, respectively. The present time step is represented by t. $\odot$ is an element-wise multiplication function (Hadamard Product) and σ is a sigmoid activation operation. i_t, f_t, g_t and o_t are the input, forget, cell and output gates in the present state, respectively. The learnable weights and bias terms between gates are represented by w and b. LSTM's fundamental innovation is the cell state s_t. The cell state, in contrast to the fast transition in the hidden state h_t, may recollect a more comprehensive history. Its memory is controlled by the forget gate f_t. The hidden state h_t is obtained from s_t using an output gate. The structure of a standard LSTM neural network is shown in Fig. 4 [31].

3.4 Dataset

The dataset is available at https://www.kaggle.com/datasets/borapajo/food-choices. The dataset includes 49 features; 48 are the predictors, and one feature is the target feature. The description of the features is demonstrated in Table 2. The violin plot for the target feature is shown in Fig. 5. The correlation matrix for the features is presented in Fig. 6. The statistical calculation for the features is demonstrated in Table 3.

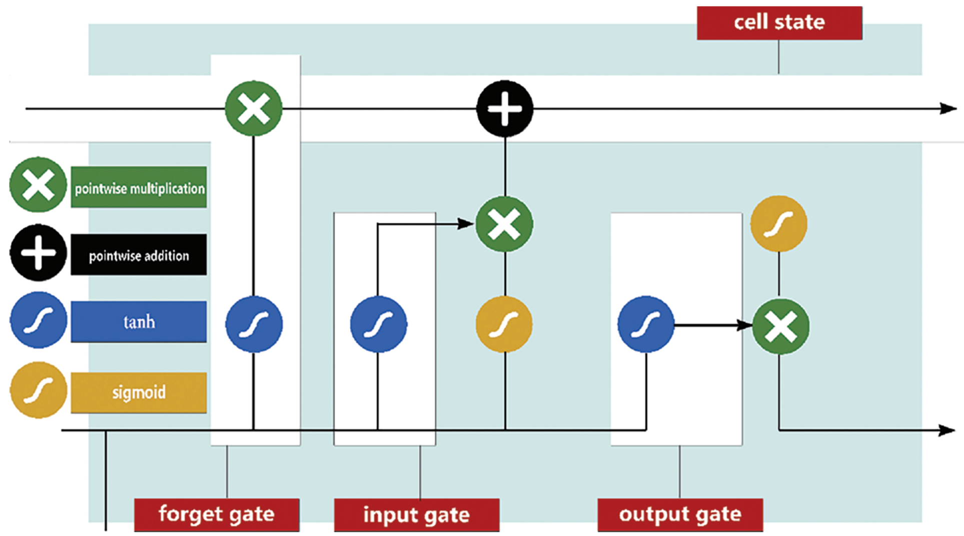

Figure 4: The structure of a standard LSTM neural network

Table 2: Description of the features

Features	Description
GPA	Grade point average (numerical value)
Gender_type	1-female and 2-male
Breakfast_type	1-cereal option and 2-donut option
Calories_of_chicken	Guessing calories in chicken piadina
No_calories_day	Importance of consuming calories per day
Guess_calories_scone	Guessing calories in a scone from starbucks
Coffee_type	1-creamy frapuccino and 2-espresso
Comfort_food_coded	It is based on the reasons that make you eat comfort food
How_cooking	How often do you cook?
Comfort_food_coded1	It is based on the reasons that make you eat comfort food
Cuisine_type	What type of cuisine did you eat growing up?
Diet_coded	It is based on words used to describe the diet
Drink_type	1-orange juice and 2-soda
Eating_coded	It is based on the eating changes
Eating_coded1	It is based on the eating changes
Eating_freq	Frequency of eating out in a typical week
Employ_type	Is the work full time or part time or no time or other
Ethnic	How likely to eat ethnic food?
Exercise_regular	How often do you exercise in a regular week?

(Continued)

Table 2 (continued)

Features	Description
Education_of_father	what is the father education?
Fav_cuisine	It is based on the favorite cuisine you prefer
Food_favorite	What is your favorite cuisine?
Fries_type	1-Mcdonald's fries or 2-home fries
Day_fruit	How likely to eat fruit in a regular day?
Grade_type	4-senior 3-junior 2-sophomore 1-freshman
Food_greek	How likely to eat greek food when available?
Feeling	It is from 1 to 10, where 1 is strongly agree and 10 is strongly disagree
Ideal_diet	Description for the ideal diet
Income	Income
Food_indian	How likely are you to eat Indian food?
Food_ italian	How likely are you to eat Italian food?
Rewarding_life	It is from 1 to 10, where 1 is strongly agree and 10 is strongly disagree
Marital_status	Marital status
Education_of_mother	what is the mother education?
Checking_ nutritional	Checking nutritional values frequency
Campus_on_off	Describes the living situation
Cooking_parents	How many days a week did the parents cook?
Meal_pay	How much would you pay for meal out?
Persian_type	How likely to eat Persian food when available?
Soup	1-veggie soup and 2-creamy soup
SPW	Self-perception of weight
Sport_type	1-yes and 2-no
Food_thai	How likely to eat Thai food when available?
Calories_ tortilla	1–580 or 2–725 or 3–940 or 4–1165
Calories_ turkey	1–345 or 2–500 or 3–690 or 4–850
Days_veggies	How likely to eat veggies in a day?
Vitamins	1-yes and 2-no
Calories_ of_waffle	1–575 or 2–760 or 3–900 or 4–1315
Weights	Weights in pounds

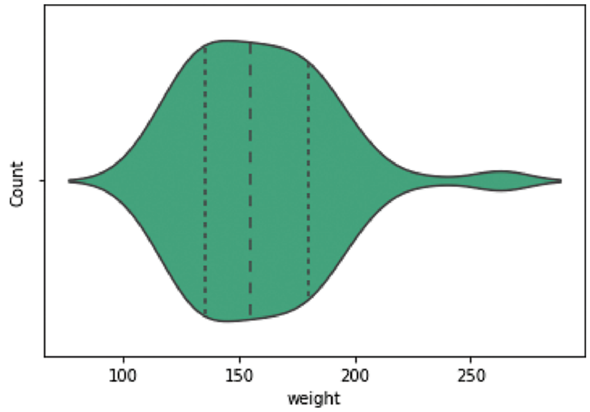

Figure 5: Violin plot for the target feature

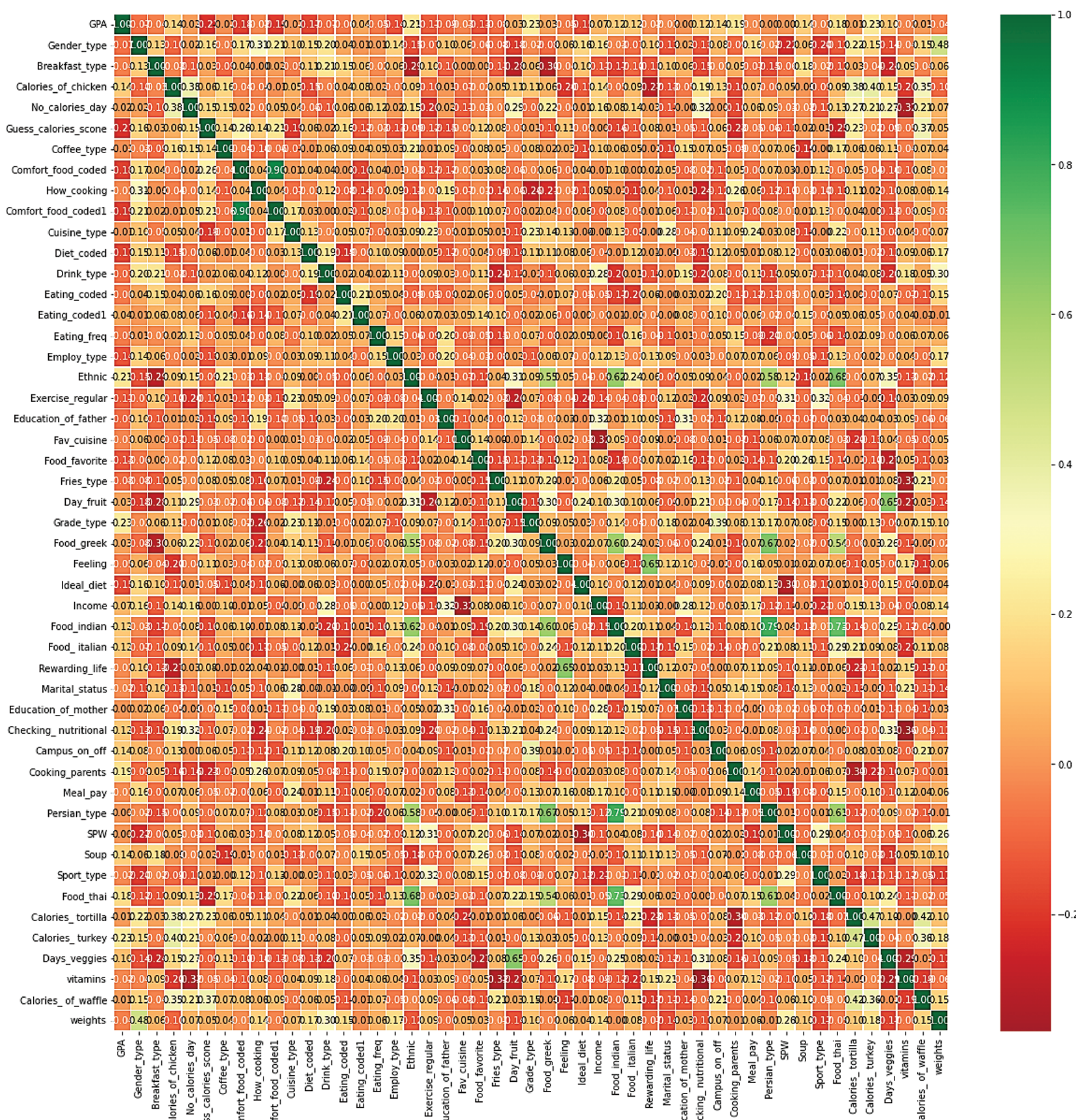

Figure 6: Correlation matrix for the features

Table 3: Statistical calculation of the features

Statistical analysis	Count	Mean	Std	Min	25%	50%	75%	Max
GPA	125.0	3.418936	0.38255	2.2	3.2	3.5	3.7	4.0
Gender_type	125.0	1.392000	0.49016	1.0	1.0	1.0	2.0	2.0
Breakfast_type	125.0	1.112000	0.31663	1.0	1.0	1.0	1.0	2.0
Calories_of_chicken	125.0	577.3200	131.214	265	430	610	720	720
No_calories_day	125.0	3.024000	0.58838	2.0	3.0	3.0	3.0	4.0
Guess_calories_scone	125.0	504.5600	230.034	315	420	420	420	980
Coffee_type	125.0	1.752000	0.43359	1.0	2.0	2.0	2.0	2.0
Comfort_food_coded	125.0	2.592000	1.83204	1.0	2.0	2.0	3.0	9.0
How_cooking	125.0	2.792000	1.02623	1.0	2.0	3.0	3.0	5.0
Comfort_food_coded1	125.0	2.688000	1.91098	1.0	2.0	2.0	3.0	9.0
Cuisine_type	125.0	1.336000	0.91531	1.0	1.0	1.0	1.0	6.0
Diet_coded	125.0	1.760000	0.76622	1.0	1.0	2.0	2.0	4.0
Drink_type	125.0	1.568000	0.49734	1.0	1.0	2.0	2.0	2.0
Eating_coded	125.0	1.536000	0.75715	1.0	1.0	1.0	2.0	4.0
Eating_coded1	125.0	4.552000	2.54778	1.0	3.0	4.0	5.0	13
Eating_freq	125.0	2.560000	1.13876	1.0	2.0	2.0	3.0	5.0
Employ_type	125.0	2.416000	0.52646	1.0	2.0	2.0	3.0	3.0
Ethnic	125.0	3.744000	1.17709	1.0	3.0	4.0	5.0	5.0
Exercise_regular	125.0	1.528000	0.65471	1.0	1.0	1.0	2.0	3.0
Education_of_father	125.0	3.488000	1.20225	1.0	2.0	4.0	4.0	5.0
Fav_cuisine	125.0	2.424000	1.94796	0.0	1.0	1.0	4.0	8.0
Food_favorite	125.0	1.704000	0.90717	1.0	1.0	1.0	3.0	3.0
Fries_type	125.0	1.088000	0.28443	1.0	1.0	1.0	1.0	2.0
Day_fruit	125.0	4.224000	0.92338	1.0	4.0	5.0	5.0	5.0
Grade_type	125.0	2.376000	1.13353	1.0	1.0	2.0	3.0	4.0
Food_greek	125.0	3.488000	1.36556	1.0	3.0	4.0	5.0	5.0
Feeling	125.0	5.456000	2.58564	1.0	3.0	5.0	8.0	10
Ideal_diet	125.0	3.704000	2.08691	1.0	2.0	3.0	6.0	8.0
Income	125.0	4.536000	1.45105	1.0	4.0	5.0	6.0	6.0
Food_indian	125.0	3.152000	1.48680	1.0	2.0	3.0	5.0	5.0
Food_ italian	125.0	4.728000	0.58717	3.0	5.0	5.0	5.0	5.0
Rewarding_life	125.0	5.104000	3.10780	1.0	2.0	5.0	8.0	10
Marital_status	125.0	1.496000	0.54807	1.0	1.0	1.0	2.0	4.0
Education_of_mother	125.0	3.440000	1.15981	1.0	2.0	4.0	4.0	5.0
Checking_ nutritional	125.0	3.152000	1.20520	1.0	2.0	3.0	4.0	5.0

(Continued)

Table 3 (continued)

Statistical analysis	Count	Mean	Std	Min	25%	50%	75%	Max
Campus_on_off	125.0	1.320000	0.67918	1.0	1.0	1.0	1.0	4.0
Cooking_parents	125.0	1.528000	0.74677	1.0	1.0	1.0	2.0	5.0
Meal_pay	125.0	3.408000	1.04028	2.0	3.0	3.0	4.0	6.0
Persian_type	125.0	2.808000	1.41817	1.0	2.0	3.0	4.0	5.0
Soup	125.0	3.120000	1.11152	1.0	2.0	3.0	4.0	6.0
SPW	125.0	1.216000	0.41317	1.0	1.0	1.0	1.0	2.0
Sport_type	125.0	1.384000	0.48831	1.0	1.0	1.0	2.0	2.0
Food_thai	125.0	3.336000	1.43652	1.0	2.0	3.0	5.0	5.0
Calories_ tortilla	125.0	947.5200	201.274	580	725	940	1165	116
Calories_ turkey	125.0	555.0400	152.370	345	500	500	690	850
Days_veggies	125.0	4.008000	1.08133	1.0	3.0	4.0	5.0	5.0
Vitamins	125.0	1.512000	0.50186	1.0	1.0	2.0	2.0	2.0
Calories_ of_waffle	125.0	1073.400	248.667	575	900	900	131	1315
Weights	125.0	158.3600	31.1190	100	135	155	180	265.0

3.5 Performance Metrics

Five error analysis criteria are presented to analyze the suggested models in order to verify the efficacy and performance of the prediction models. These criteria are given in Eqs. (10)–(14), where y_{real_i} is the actual values, y_{pred_i} is the predicted values, N is the sample size, and $\bar{y}$ is the mean of the actual values [32,33]. The effectiveness of each model is assessed using the median absolute error (MedAE), mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE) and mean absolute percentage error (MAPE).

$$\mathrm{MAE} = \frac{1}{N}\sum\nolimits_{i=1}^{N}\left|y_{real_i} - y_{pred_i}\right| \tag{10}$$

$$\mathrm{MedAE} = median\left(\left|y_{real_1} - y_{pred_1}\right|,\ \ldots\ldots,\ \left|y_{real_N} - y_{pred_N}\right|\right) \tag{11}$$

$$\mathrm{MSE} = \frac{1}{N}\sum\nolimits_{i=1}^{N}\left(y_{real_i} - y_{pred_i}\right)^2 \tag{12}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum\nolimits_{i=1}^{N}\left(y_{real_i} - y_{pred_i}\right)^2} \tag{13}$$

$$\mathrm{MAPE} = \left(\frac{1}{N}\sum\nolimits_{i=1}^{N}\left|\frac{y_{real_i} - y_{pred_i}}{y_{real_i}}\right|\right) \times 100 \tag{14}$$

4 Results and Discussions

The major experiments used to assess the effectiveness of the stacked-LSTM model are detailed in this section. A version of the Jupyter notebook (6.4.6) was used to execute the stacked-LSTM model. The Jupyter notebook simplifies the process of running and writing Python code. It is a popular open-source tool for building and running deep learning and machine learning algorithms for classification and regression. In this study, the performance of the stacked-LSTM model was compared with the models, namely, LSTM, CNN-LSTM, CNN, and KNN. The performance of these models was evaluated using MAE, MedAE, MSE, MAPE, and RMSE as the evaluation metrics. In the LSTM model, the number of hidden units is 100. The batch size is 32, the used learning rate is 0.001, the iterations number is 100, the Adam optimizer is the used optimizer, the time steps are 48, and the activation function in the output is a linear function. In the stacked-LSTM model, three layers are used. The first layer consists of 128 hidden units, the second layer consists of 128 hidden units, and the third layer consists of 32 hidden units. The batch size is 32, the used learning rate is 0.001, the iterations number is 100, the used optimizer is the Adam optimizer, the time steps are 48, and the used activation function is a linear function. In the CNN-LSTM model, the model includes three convolution layers, one max-pooling layer, one LSTM layer, one hidden layer, and an output layer that returns a single, continuous value. The first convolution layer consists of 64 filters. The number of kernel sizes is 7. The second convolution layer contains 32 filters. The number of kernel sizes is 3. The third convolution layer contains 16 filters. The number of kernel sizes is 3. The LSTM layer includes 128 hidden units. The hidden layer includes 32 neurons. The activation function used in the output layer is the linear function. In the CNN model, the model consists of three convolution layers one max-pooling layer, one hidden layer, and an output layer that returns a single, continuous value. The first convolution layer contains 64 filters. The number of kernel sizes is 7. The second convolution layer contains 32 filters. The number of kernel sizes is 3. The third convolution layer contains 16 filters. The number of kernel sizes is 3. The hidden layer includes 32 neurons. The activation function used in the output layer is the linear function. In the KNN Regressor model, the parameters used for the model are presented in Table 4. Five evaluation metrics, namely, mean squared error (MSE), mean absolute error (MAE), median absolute error (MedAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE), were computed for the models, namely, the LSTM model, stacked-LSTM model, CNN-LSTM model, CNN model, and KNN regressor model, respectively. Table 5 demonstrates the performance of the models utilized in this study.

Table 4: Parameters used for the KNN regressor model

Model	Parameters
KNN regressor	n_neighbors = 1, weights = distance.

Table 5: Performance of the models

Model	MAE	MSE	MedAE	RMSE	MAPE
LSTM	0.0100	0.00020	0.009	0.0143	22.4
Stacked-LSTM	**0.0087**	**0.00011**	**0.006**	**0.011**	**21.0**
CNN-LSTM	0.0093	0.00018	0.0085	0.017	21.7
CNN	0.0099	0.00014	0.0082	0.015	21.5
KNN regressor	0.0093	0.00012	0.0075	0.014	21.2

As shown in Table 5, the stacked-LSTM model archives better results. Its MAE, MSE, MedAE, RMSE, and MAPE are 0.0087, 0.00011, 0.006, 0.011, and 21, respectively. The LSTM model achieved the lowest results among all models; its MAE, MSE, MedAE, RMSE, and MAPE were 0.0100, 0.00020, 0.009, 0.0143, and 22.4, respectively. Figs. 7–10 demonstrate the relationship between mean absolute error, mean squared error and number of epochs using the LSTM model, stacked-LSTM model, CNN-LSTM model, and CNN model, respectively. Fig. 11 illustrates a comparison between the actual error rate values and the predicted error values for the models, namely, LSTM, stacked-LSTM model, CNN model, CNN-LSTM model, and KNN model, respectively.

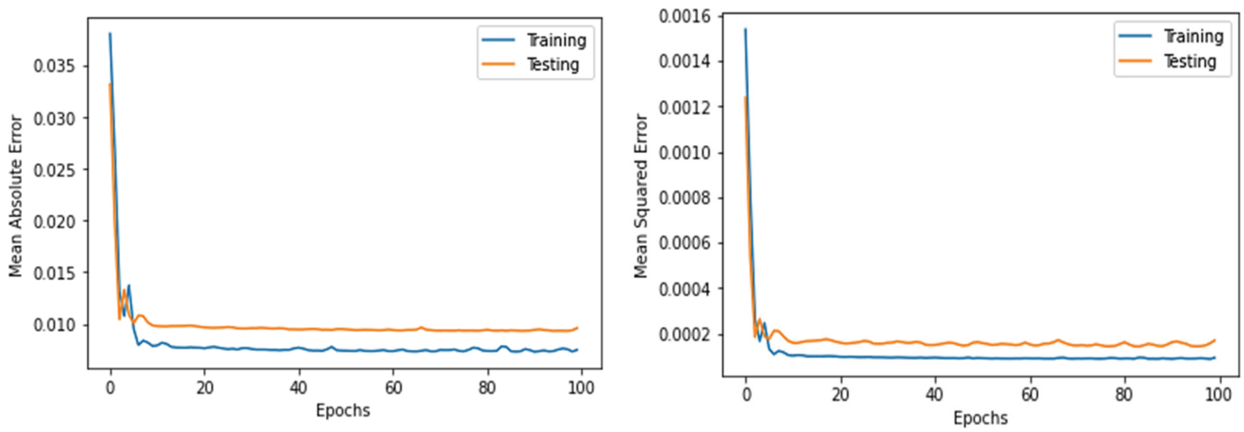

Figure 7: Relation between mean absolute error, mean squared error and number of epochs using LSTM

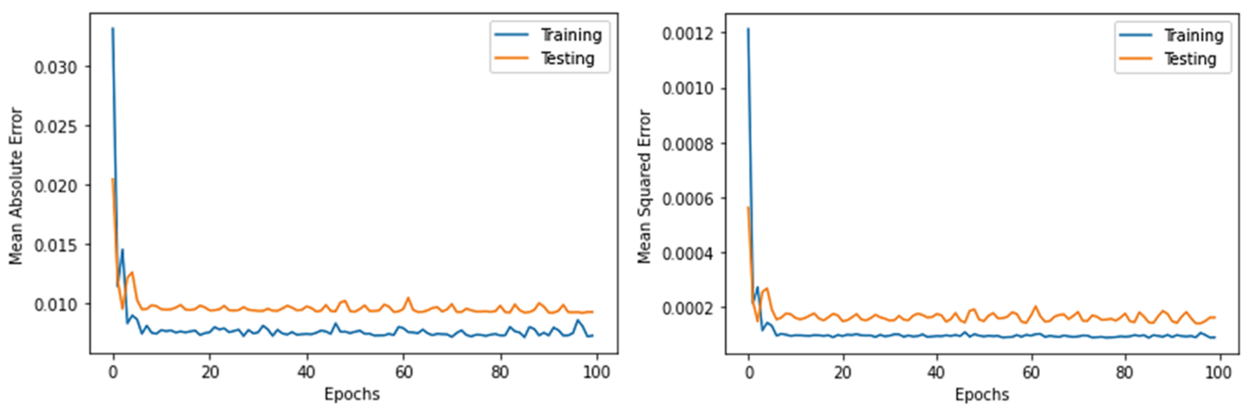

Figure 8: Relation between mean absolute error, mean squared error and number of epochs of stacked-LSTM

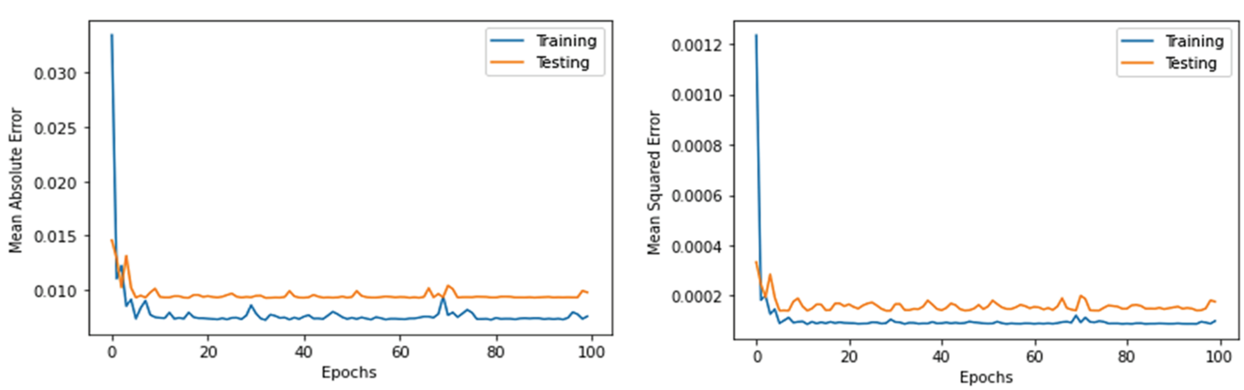

Figure 9: Relation between mean absolute error, mean squared error and number of epochs of CNN-LSTM

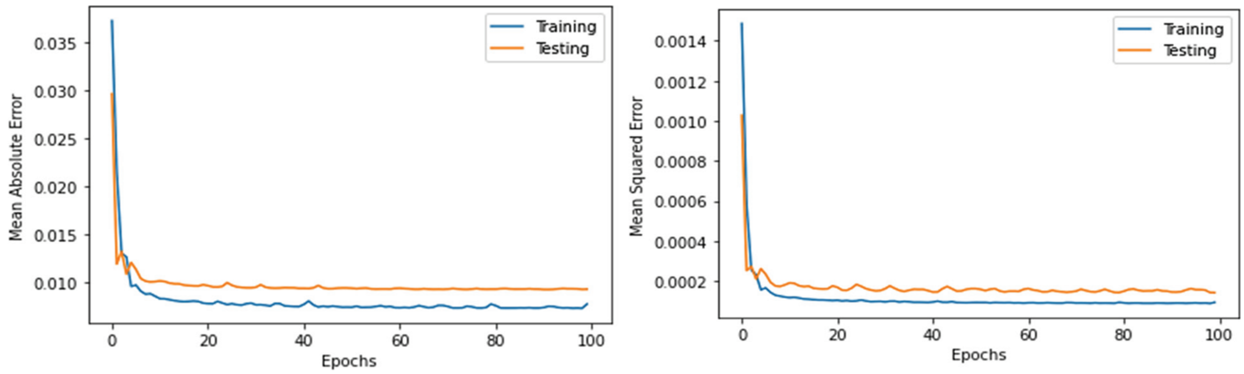

Figure 10: Relation between mean absolute error, mean squared error and number of epochs using CNN model

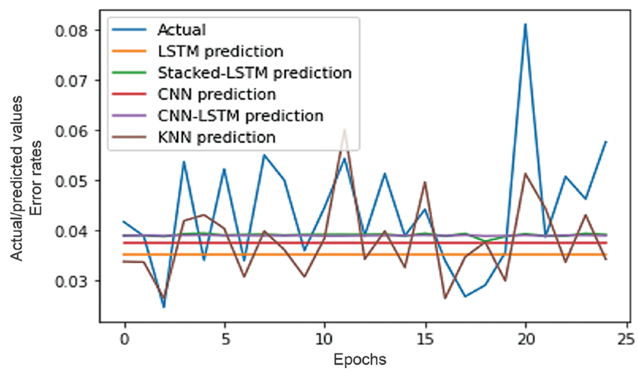

Figure 11: Comparison between actual values and predicted values

5 Conclusion and Perspectives

In this paper, an effective stacked-LSTM model was developed for the weight prediction of food choices. The proposed stacked-LSTM model uses three layers, where the first layer consists of 128 hidden units, the second layer consists of 128 hidden units, and the third layer consists of 32 hidden units. The Adam optimizer is also used. Several performance metrics, namely, MAE, MSE, MedAE, RMSE, and MAPE, are utilized to assess the suggested model's effect. The suggested model's performance was compared to four other models, namely, the LSTM model, the CNN model, the CNN-LSTM model, and KNN. The stacked-LSTM model achieved the best results among the four models, where MAE, MSE, MedAE, RMSE, and MAPE were 0.0087, 0.00011, 0.006, 0.011, and 21, respectively. Possible future work may consider augmenting the model with a food recommendation system. The effect of weight prediction in general healthcare systems as well as the impact of nutrition during the current COVID-19 epidemic are other important considerations.

Funding Statement: The authors received no specific funding for this study.

Availability of Data and Materials: A data availability is found in https://www.kaggle.com/datasets/borapajo/food-choices.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. C. Onwezen, M. J. Reinders, M. C. D. Verain and H. M. Snoek, "The development of a single-item food choice questionnaire," *Food Quality and Preference*, vol. 71, pp. 34–45, 2019.

[2] M. C. D. Verain, H. M. Snoek, M. C. Onwezen, M. J. Reinders and E. P. Bouwman, "Sustainable food choice motives: The development and cross-country validation of the sustainable food choice questionnaire (SUS-FCQ)," *Food Quality and Preference*, vol. 93, no. 104267, pp. 1–11, 2021.

[3] J. P. H. Wilding, R. L. Batterham, S. Calanna, M. Davies, L. F. Van-Gaal *et al.,* "Once-weekly semaglutide in adults with overweight or obesity," *New England Journal of Medicine*, vol. 384, no. 11, pp. 989–1002, 2021.

[4] S. Basu, "The transitional dynamics of caloric ecosystems: Changes in the food supply around the world," *Critical Public Health*, vol. 25, no. 3, pp. 248–264, 2015.

[5] J. Dunstan, M. Aguirre, M. Bastías, C. Nau, T. A. Glass *et al.,* "Predicting nationwide obesity from food sales using machine learning," *Health Informatics Journal*, vol. 26, no. 1, pp. 652–663, 2020.

[6] L. Zhu, P. Spachos, E. Pensini and K. N. Plataniotis, "Deep learning and machine vision for food processing: A survey," *Current Research in Food Science*, vol. 4, pp. 233–249, 2021.

[7] P. -J. Chen and M. Antonelli, "Conceptual models of food choice: Influential factors related to foods, individual differences, and society," *Foods*, vol. 9, no. 12, pp. 1–21, 2020.

[8] K. Bi, T. Qiu and Y. Huang, "A deep learning method for yogurt preferences prediction using sensory attributes," *Processes*, vol. 8, no. 5, pp. 1–12, 2020.

[9] M. Pellegrini, V. Ponzo, R. Rosato, E. Scumaci, I. Goitre *et al.,* "Changes in weight and nutritional habits in adults with obesity during the 'lockdown' period caused by the COVID-19 virus emergency," *Nutrients*, vol. 12, no. 7, pp. 1–11, 2020.

[10] L. J. Hintze, É. Doucet and G. S. Goldfield, "The relative reinforcing value of snack food is a significant predictor of fat loss in women with overweight or obesity," *Applied Physiology, Nutrition, and Metabolism*, vol. 47, no. 2, pp. 134–140, 2022.

[11] A. Gere, K. Héberger and S. Kovács, "How to predict choice using eye-movements data?" *Food Reseach Intenational*, vol. 143, no. 110309, pp. 1–7, 2021.

[12] J. Depa, J. R. Barrada and M. Roncero, "Are the motives for food choices different in orthorexia nervosa and healthy orthorexia?" *Nutrients*, vol. 11, no. 3, pp. 1–14, 2019.

[13] A. Nogales, R. Díaz-Morón and Á. J. García-Tejedor, "A comparison of neural and non-neural machine learning models for food safety risk prediction with european union RASFF data," *Food Control*, vol. 134, no. 108697, pp. 1–11, 2022.

[14] C. Zhu, T. Beatty, Q. Zhao, W. Si and Q. Chen, "Using genetic information to improve the prediction of individual food choice: A case study of alcoholic beverages," *bioRxiv*, vol. 44, no. 18, pp. 2–34, 2021.

[15] M. D. Amin, S. Badruddoza and J. J. McCluskey, "Predicting access to healthful food retailers with machine learning," *Food Policy*, vol. 99, no. 101985, pp. 1–22, 2021.

[16] M. Nilashi, H. Ahmadi, G. Arji, K. O. Alsalem, S. Samad *et al.,* "Big social data and customer decision making in vegetarian restaurants: A combined machine learning method," *Journal of Retailing and Consumer Services*, vol. 62, no. 102630, pp. 1–18, 2021.

[17] D. Cosme, D. Zeithamova, E. Stice and E. T. Berkman, "Multivariate neural signatures for health neuroscience: Assessing spontaneous regulation during food choice," *Social Cognitive and Affective Neuroscience*, vol. 15, no. 10, pp. 1120–1134, 2020.

[18] W. Hidayat, M. Ardiansyah and K. Kusrini, "Decision support system for selection of staples food and food commodity price prediction post-covid-19 using simple additive weighting and multiple linear regression methods," in *2020 3rd Int. Conf. on Information and Communications Technology (ICOIACT)*, IEEE, Yogyakarta, Indonesia, pp. 45–50, 2020.

[19] G. Baryannis, S. Dani and G. Antoniou, "Predicting supply chain risks using machine learning: The trade-off between performance and interpretability," *Future Generation Computer Systems*, vol. 101, pp. 993–1004, 2019.

[20] C. Bian, H. Shi, S. Wu, K. Zhang, M. Wei *et al.,* "Prediction of field-scale wheat yield using machine learning method and multi-spectral uav data," *Remote Sensing*, vol. 14, no. 6, pp. 1–18, 2022.

[21] M. Y. Shams, O. M. Elzeki, L. M. Abouelmagd, A. E. Hassanien, M. Abd Elfattah *et al.,* "HANA: A healthy artificial nutrition analysis model during COVID-19 pandemic," *Computers in Biology and Medicine*, vol. 135, no. 104606, pp. 1–16, 2021.

[22] J. Peacock and P. Peacock, "Oxford handbook of medical statistics," in *Oxford University Press*, 1st ed., vol. 1. New York, USA: Oxford University Press Inc., pp. 76–98, 2011.

[23] A. L. Samuel, "Some studies in machine learning using the game of checkers," *IBM Journal of Research and Development*, vol. 44, no. 1.2, pp. 206–226, 2000.

[24] L. Zhou, C. Zhang, F. Liu, Z. Qiu and Y. He, "Application of deep learning in food: A review," *Comprehensive Reviews in Food Science and Food Safety*, vol. 18, no. 6, pp. 1793–1811, 2019.

[25] M. A. El-Dosuky, M. Z. Rashad, T. T. Hamza and A. H. El-Bassiouny, "Food recommendation using ontology and heuristics," in *Int. Conf. in Advanced Machine Learning Technologies and Applications (AMLTA)*, Springer, Cairo, Egypt, pp. 423–429, 2012.

[26] M. Elhoseny, Z. Tarek, I. El-Hasnony, "Advanced cognitive algorithm for biomedical data processing: COVID-19 pattern recognition as a case study," *Journal of Healthcare Engineering*, vol. 2022, pp. 1–11, 2022.

[27] T. U. Rehman, M. S. Mahmud, Y. K. Chang, J. Jin and J. Shin, "Current and future applications of statistical machine learning algorithms for agricultural machine vision systems," *Computers and Electronics in Agriculture*, vol. 156, pp. 585–605, 2019.

[28] A. Y. Abyaneh, A. H. G. Foumani and V. Pourahmadi, "Deep neural networks meet csi-based authentication," *ArXiv Preprint*, ArXiv1812.04715, pp. 2–8, 2018.

[29] D. Arpit, B. Kanuparthi, G. Kerg, N. R. Ke, I. Mitliagkas *et al.,* "h-detach: Modifying the LSTM gradient towards better optimization," *ArXiv Preprint*, ArXiv1810.03023, ICLR, pp. 1–19, 2018.

[30] E. -S. M. El-Kenawy, M. M. Eid, M. Saber and A. Ibrahim, "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection," *IEEE Access*, vol. 8, pp. 107635–107649, 2020.

[31] R. H. Assaad and S. Fayek, "Predicting the price of crude oil and its fluctuations using computational econometrics: Deep learning, LSTM, and convolutional neural networks," *Econometric Research in Finance*, vol. 6, no. 2, pp. 119–137, 2021.

[32] A. Botchkarev, "Evaluating performance of regression machine learning models using multiple error metrics in azure machine learning studio," SSRN3177507, pp. 1–16, 2018. [Online]. Available: https://ssrn.com/abstract=3177507.

[33] M. Y. Shams, A. Tolba and S. Sarhan, "A vision system for multi-view face recognition," *International Journal of Circuits, Systems, and Signal Processing*, vol. 10, no. 1, pp. 455–461, 2017.