

METHODOLOGY

Open Access



A deep learning approach for classifying and predicting children's nutritional status in Ethiopia using LSTM-FC neural networks

Getnet Bogale Begashaw^{1,2*}, Temesgen Zewotir³ and Haile Mekonnen Fenta^{1,4,5}

*Correspondence:

Getnetbogale145@gmail.com;
Getnetbogale@dbu.edu.et

¹ Department of Statistics, College of Science, Bahir Dar University, P.O. Box 79, Bahir Dar, Ethiopia

² Department of Data Science, College of Natural and Computational Science, Debre Berhan University, P.O. Box 445, Debre Berhan, Ethiopia

³ School of Mathematics, Statistics and Computer Science, College of Agriculture, Engineering and Science, University of KwaZulu-Natal, Durban, South Africa

⁴ Center for Environmental and Respiratory Health Research, Population Health, University of Oulu, Oulu, Finland

⁵ Biocenter Oulu, University of Oulu, Oulu, Finland

Abstract

Background: This study employs a LSTM-FC neural networks to address the critical public health issue of child undernutrition in Ethiopia. By employing this method, the study aims classify children's nutritional status and predict transitions between different undernutrition states over time. This analysis is based on longitudinal data extracted from the Young Lives cohort study, which tracked 1,997 Ethiopian children across five survey rounds conducted from 2002 to 2016. This paper applies rigorous data preprocessing, including handling missing values, normalization, and balancing, to ensure optimal model performance. Feature selection was performed using SHapley Additive exPlanations to identify key factors influencing nutritional status predictions. Hyperparameter tuning was thoroughly applied during model training to optimize performance. Furthermore, this paper compares the performance of LSTM-FC with existing baseline models to demonstrate its superiority. We used Python's TensorFlow and Keras libraries on a GPU-equipped system for model training.

Results: LSTM-FC demonstrated superior predictive accuracy and long-term forecasting compared to baseline models for assessing child nutritional status. The classification and prediction performance of the model showed high accuracy rates above 93%, with perfect predictions for Normal (N) and Stunted & Wasted (SW) categories, minimal errors in most other nutritional statuses, and slight over- or underestimations in a few instances. The LSTM-FC model demonstrates strong generalization performance across multiple folds, with high recall and consistent F1-scores, indicating its robustness in predicting nutritional status. We analyzed the prevalence of children's nutritional status during their transition from late adolescence to early adulthood. The results show a notable decline in normal nutritional status among males, decreasing from 58.3% at age 5 to 33.5% by age 25. At the same time, the risk of severe undernutrition, including conditions of being underweight, stunted, and wasted (USW), increased from 1.3% to 9.4%.

Conclusions: The LSTM-FC model outperforms baseline methods in classifying and predicting Ethiopian children's nutritional statuses. The findings reveal a critical rise in undernutrition, emphasizing the need for urgent public health interventions.

Keywords: LSTM-FC, Classification, Feature selection, Prediction, Young lives cohort study



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Child undernutrition is a condition characterized by an inadequate intake of essential nutrients, leading to stunted growth, weakened immune systems, and developmental delays [1–3]. This condition is commonly assessed using the World Health Organization (WHO) growth standards, which categorize nutritional deficiencies into three main types: underweight (low weight-for-age), stunting (low height-for-age), and wasting (low weight-for-height) [4–11].

Globally, undernutrition significantly impacts young children, with approximately 149 million children under five years old being stunted and 49.5 million wasted, accounting for over 45% of the disease burden in this age group [12, 13]. In Asia, there are 78 million stunted children, including 31 million in India alone. As of 2020, the number of stunted children in Asia was 68 million, closely followed by Africa with 64 million [14]. It was estimated that 21.3% (144 million) of all children under five were stunted, with 36% residing in Sub-Saharan Africa (SSA) and Southern Asia, regions that also account for nearly 90% of underweight children [15–17]. Although the prevalence of stunting in SSA decreased from 34.5% in 2012 to 31.1% in 2019, the reduction has not been swift enough to meet global targets [16]. Furthermore, child deaths from undernutrition are nearly double those from diarrhea and pneumonia combined, highlighting the severe impact of this issue [18].

Given this critical global health challenge, requiring ongoing and innovative research methods to fully understand its causes, trends, and future trajectories. In this context, the recurrent neural network (RNN), a recent advancement in deep learning, emerges as a powerful tool for analyzing and predicting patterns in time-series data [19, 20]. Deep learning techniques, including RNNs, have made significant contributions across various domains such as healthcare, finance, meteorology, autonomous driving, natural language processing, and technologies for recognizing images, voices, and handwriting [21–26]. A specialized type of RNN, Long Short-Term Memory (LSTM) networks, was introduced by [27] to address the limitations of traditional RNNs, particularly their short memory span and the challenges of vanishing and exploding gradients [28, 29].

The prevalence and related factors of child undernutrition have been studied using various statistical approaches, including regression [30–37], longitudinal models [38, 39], Markov chain [40, 41], Bayesian network [42, 43], and machine learning techniques [44–50]. However, these traditional statistical methods have limitations in managing long-term dependencies and temporal patterns in sequential data. For instance, Markov chain models assume memorylessness, predicting future states based solely on the current state, which restricts their ability to track and learn from entire data sequences [51, 52]. Similarly, Dynamic Bayesian Networks are limited by their predefined structures and the conditional independence assumptions of each variable, hindering their capacity to model complex, non-linear relationships in data [53, 54]. Traditional machine learning methods, such as decision trees and support vector machines, also struggle to capture temporal dynamics and may inadequately handle sequential data.

In contrast, deep learning models like LSTMs effectively address these challenges through memory cells and gating mechanisms [55–60]. LSTMs excel in managing sequential data, capturing complex patterns, and enhancing accuracy in predicting children's future nutritional status [19, 61–63]. Building on this capability, LSTM models are often

used to forecast various diseases, such as Parkinson's, bronchopneumonia, dementia, and Influenza-Like Illness [64–67]. Their success in capturing complex temporal patterns and predicting disease trends highlights their potential. However, the application of LSTM models for child undernutrition remains limited.

Many studies on child undernutrition have employed composite indices [45, 50, 68] or categorized children into groups such as normal, underweight, stunted, and wasted, or by severity as mild, moderate, and severe [44, 46, 69–75]. For example, machine learning models conducted in countries such as Bangladesh [26, 50, 73, 74], India [47, 76–78], Nigeria [74], Ghana [46], and Ethiopia [44, 45, 79] often utilized these approaches. However, these studies did not account for the possibility of concurrent nutritional outcomes, where children might exhibit multiple forms of undernutrition simultaneously. To address this limitation, we employed a more nuanced deep learning classifier approach that considers the probability of children being in multiple states of undernutrition by considering gender difference.

Deep learning classifiers have been applied to early detection of children's nutritional status, as evidenced by studies from Bangladesh [26], Peru [80], India [81–84], and USA [85]. Most of these studies primarily highlight the effectiveness of deep learning algorithms using images, rather than longitudinal follow-up data. However, despite Ethiopia having one of the highest rates of undernourished children, the application of deep learning algorithms in this context remains relatively underexplored.

Moreover, extensive research has focused on undernutrition among children under five, emphasizing its critical impact during early childhood [4, 5, 32, 36–38, 45]. However, studies beyond this stage remain limited, even though early undernutrition often persists, resulting in long-term health and developmental challenges.

As far as we are aware, there is limited evidence regarding the use of deep learning classifier approach in comparison with other baseline methods to assess concurrent nutritional statuses in children from early childhood through adolescence in Ethiopia. Consequently, this proof-of-concept study seeks to explore the potential of an LSTM-FC neural network against models like Auto-Regressive Integrated Moving Average (ARIMA), Random Forest (RF), Recurrent Neural Network (RNN), and Gated Recurrent Unit (GRU), and Bi-LSTM in classifying and predicting undernutrition among children under 15. Our study forecasts the future nutritional status of children over the age of 15, identifying which status is more prevalent and which is less common. As children gain independence and become less influenced by caregivers and families, identifying these trends in nutritional status becomes increasingly important.

Finally, this paper is organized as follows: Sect. 1 introduces the study background, motivation, and objectives. Section 2 details the methodology, including data preprocessing, model development, and evaluation metrics. Section 3 presents the results and analysis. Section 4 discusses the findings in the context of existing literature. Finally, Sect. 5 concludes with key insights, limitations, and recommendations for future research.

Methods

Data source and study design

This study utilized longitudinal data from Ethiopia's Young Lives of Young cohort (YLCS), an international initiative aimed at addressing childhood poverty and health. The cohort includes approximately 1999 children aged 1–15 [86, 87]. The country is

highly heterogeneous, with large socioeconomic differences across regional states and between urban and rural areas [88]. The surveys were conducted at 20 sentinel sites across five Ethiopian regions: Amhara, Oromia, Southern Nations, Nationalities, and Peoples (SNNP), Tigray, and Addis Ababa City Administration (CA). Data collection occurred in five distinct waves—2002, 2006, 2009, 2013, and 2016—each spaced at approximately equal intervals throughout the study period [89]. The original dataset consists of 9,995 observations, representing 1,999 children, with each child having 5 visits, including a zero placeholder for any missed visits. The training data initially includes 8,556 observations, but after applying SMOTE for data augmentation, the training set expanded to 30,208 observations, while the testing data remains at 1,440 observations.

Intervention programs were available across the study regions, targeting specific vulnerable populations. The Productive Safety Net Program (PSNP) operated in 14 sentinel sites across four regions (excluding Addis Ababa City Administration), focusing on pro-poor households [90]. Meanwhile, the Emergency Aid Program (EAP) and Health Extension Program (HEP) were available in all five regions, addressing the needs of disadvantaged socioeconomic groups by offering antenatal care, childhood disease management, and micronutrient supplementation [91].

To identify beneficiaries, the study conducted interviews with randomly selected households to assess their participation in the PSNP, EAP, and/or HEP within the past 12 months. Households were categorized as beneficiaries or non-beneficiaries for PSNP and EAP beginning in 2009 (third wave), while HEP categorization started in 2013 (fourth wave). These intervention programs were later consolidated into a single variable, “program participation status,” encompassing eight categories (C, P, E, H, PE, PH, EH, and PEH), as detailed in Table 1 at [41].

Variables in the study

Outcome variable

The YLCS utilized WHO standards to standardize Z scores for each observation, defining children’s anthropometric status based on categories such as underweight, stunted, or wasted. After cleaning and making necessary adjustments to the dataset, we classified each child’s anthropometric condition at each time point: U for underweight only, S for stunted only, W for wasted only, US for underweight and stunted, UW for underweight and wasted, SW for stunted and wasted, and USW for all three conditions. Furthermore, the condition is classified as Normal (N) if the child is not underweight, stunted, or wasted, indicating their measurements are within the expected range for their age and sex [92]. The classification percentage of children’s nutritional status over time are presented in Table S2.

Table 1 Performance metrics of LSTM-FC model across different time-based data splits

Train set	Test set	Sensitivity	Specificity	Precision	F1	AUC (95% CI)	Accuracy (95% CI)
2002 and 2006	2009	0.72	0.68	0.7	0.71	0.74 (0.70–0.78)	0.69 (0.65–0.73)
2002, 2006, and 2009	20,013	0.78	0.75	0.76	0.77	0.80 (0.76–0.84)	0.77 (0.73–0.81)
2002, 2006, 2009, and 2013	2016	0.85	0.84	0.83	0.84	0.90 (0.87–0.93)	0.83 (0.79–0.87)
2006 and 2009	20,013	0.75	0.72	0.74	0.74	0.78 (0.74–0.82)	0.73 (0.69–0.77)
2009 and 2013	2016	0.81	0.79	0.8	0.8	0.84 (0.80–0.88)	0.82 (0.78–0.86)

Features

The list of features (independent variables) included in this study were extracted from the dataset and have been extensively discussed in the literature [93–97]. These variables serve as an input feature (x) to the LSTM-FC neural network, which processes them internally through its gates to model temporal dependencies and make predictions. A list of these independent variables, along with their distribution, bivariate analysis and corresponding visualizations, is provided in the supplementary material.

Data preprocessing pipeline

The research workflow starting with data preprocessing, which involves handling missing values and identifying outliers to ensure analysis reliability before proceeding to feature selection (Fig. 1). To handle missing values, we applied mean imputation for continuous variables and mode imputation for categorical ones, repeating the process until all variables were fully complete [98]. Furthermore, string data was converted to numeric values using Python’s widely-used ‘Numpy’ library. The categorical data was converted using one-hot encoding to make it compatible with deep learning algorithms in the LSTM-FC model. Similarly, as detailed in the supplementary material, Min–Max normalization was applied to scale the features. This process enhanced data consistency, leading to improved model performance and increased prediction accuracy.

Furthermore, we employed the Synthetic Minority Oversampling Technique (SMOTE) to address class imbalance in the dataset, as outlined in the supplementary material. SMOTE generates synthetic samples for minority classes, promoting a more balanced class distribution. This technique enhances model performance and prediction accuracy, particularly for minority classes, by reducing bias and improving the robustness of the results. Initially, the training dataset consisted of 7,582 samples, with 3,706 instances representing the majority class (normal) and only 6 instances for the minority

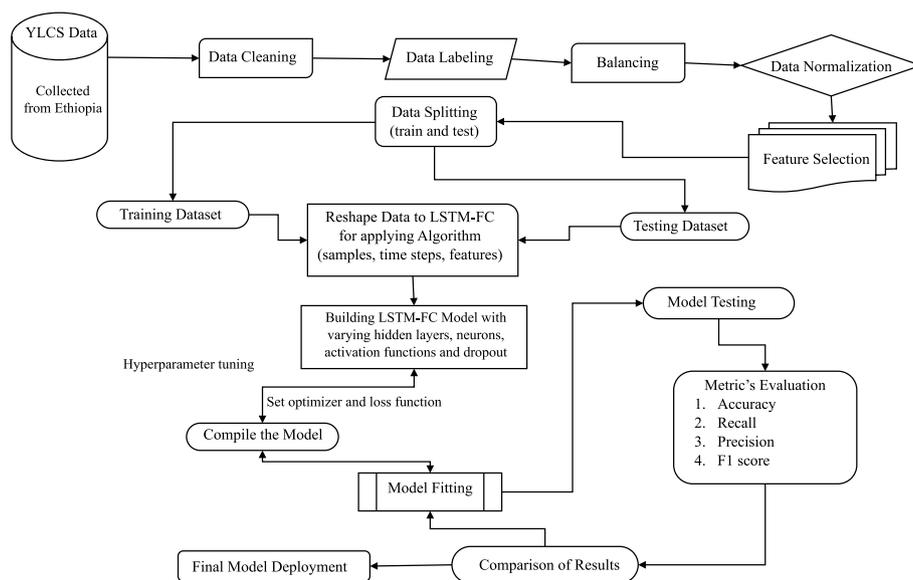


Fig. 1 Workflow Diagram of the LSTM-FC Model for Predicting Child Undernutrition in Ethiopia (2002–2016) [26, 99]

class (stunting and wasting concurrently). After applying SMOTE, the training dataset was expanded to 29,648 samples (Fig. S3). This augmentation not only balanced the class distribution but also improved the model’s ability to learn from the minority class.

For feature selection, we employed SHAP (Shapley Additive Explanations), a method that helps to identify the most impactful features by providing insights into the contribution of each feature to the model’s predictions. This approach enhances interpretability and allows for a more informed selection of features, ensuring that the model focuses on the most relevant variables. Features with higher absolute SHAP values are considered more influential in the model’s decision-making process. This approach aligns with established practices for feature selection in LSTM models, enhancing the reliability and robustness of our results [85, 100, 101]. As shown in Fig. 2, the mean absolute SHAP values for features above the median are presented. Due to publication length constraints, SHAP values for all variables are provided in supplementary material.

In this study, a time-based data splitting strategy was employed to account for the longitudinal nature of the dataset and ensure temporal consistency between training and testing, and align with the requirements of the LSTM-FC model. A recommended approach for longitudinal data is to train the model using earlier time points and test it on the latest data, as this reflects real-world forecasting scenarios where historical data is used to predict future outcomes. This approach ensures the model learns temporal dynamics over time and is evaluated on its ability to generalize to unseen, future data.

Architecture of LSTM-FC neural network

In this study, we used an LSTM-FC model to analyze sequences of health and socio-demographic data over time to predict trends in child undernutrition. As illustrated in Fig. 2, LSTM-FC networks employ several key gates to manage the flow of information.

- a Forget Gate (f_t): The forget gate determines which parts of the previous cell state (c_{t-1}) to retain or discard, filtering out unnecessary historical information about a child’s nutritional status for predicting future outcomes. The forget gate is computed using the formula [102]:

$$f_t = \sigma(\mathbf{W}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \tag{1}$$

where σ is the sigmoid activation function, which compresses the output between 0 and 1; \mathbf{W}_f is the weight matrix for the forget gate; \mathbf{b}_f is the bias term; \mathbf{h}_{t-1} represents the hidden state from the previous time step; and \mathbf{x}_t is the current input.

- b Input Gate (i_t): This gate determines which new information from the current input or features (\mathbf{x}_t), and the previous hidden state (\mathbf{h}_{t-1}) should be added to the cell state. It manages the updating process by regulating the incorporation of new features into the cell state. The input gate can be calculated using [103]:

$$i_t = \sigma(\mathbf{W}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) \tag{2}$$

where \mathbf{W}_i is the weight matrix for the input gate, and \mathbf{b}_i is the corresponding bias term.

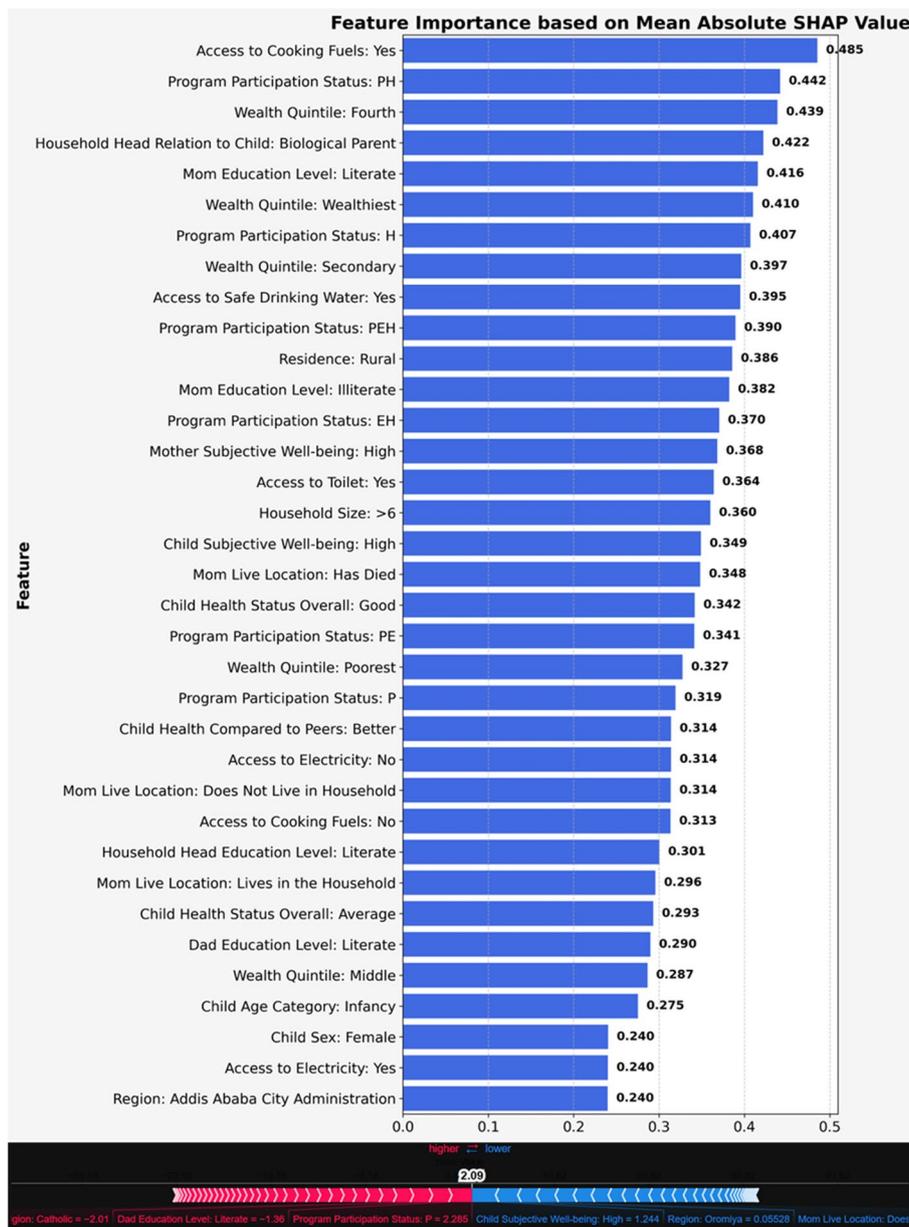


Fig. 2 Mean Absolute SHAP Value (average impact on model output) for features above the median). Note: Mean absolute SHAP values for all features are provided in the supplementary material due to publication length constraints

- c Candidate Cell State (\tilde{c}_t): This component generates a vector of potential new values that could be integrated into the cell state, representing potential changes in the child’s nutritional status based on new data inputs. Combined with the input gate, it contributes to updating the cell state. The candidate cell state is given by [103]:

$$\tilde{c}_t = \tanh(\mathbf{W}_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

where \mathbf{W}_c is the weight matrix for the candidate cell state; b_c is the bias term; and \tanh is the hyperbolic tangent activation function, which scales values between -1 and 1 .

- d Cell State Update (c_t): The cell state (c_{t-1}) is updated by merging the previous cell state (c_{t-1}), forget gate output (f_t), and the candidate cell state (\tilde{c}_t). This update reflects the current understanding of the child's nutritional status, combining past trends and new information. It serves as a memory that helps retain information over long sequences. The update is performed as [102]:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{4}$$

Here, \odot denotes element-wise multiplication, which means that two vectors or matrices of the same dimensions are multiplied together element by element.

- e Output Gate (o_t): This gate determines which elements of the updated nutritional state should be used to make predictions about the child's nutritional status in the next time step. It determines what information from the cell state will be passed on to the next layer or time step. It can be computed as [103]:

$$O_t = \sigma (W_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

where W_o is the weight matrix for the output gate, and b_o is the bias term.

- f Hidden State (h_t): The final output of the LSTM cell is produced, which could represent the predicted likelihood of undernutrition or the expected nutritional status of the child at the next time step, based on the current cell state (c_t) and the output gate (O_t). The hidden state is calculated as [102]:

$$h_t = o_t \odot \tanh(c_t) \tag{6}$$

where h_t represents the hidden state at time t , $\tanh(c_t)$ is the hyperbolic tangent of the cell state at time t , and o_t is the output gate at time t , and the symbol \odot denotes element-wise multiplication, combining the output gate with the tanh of the cell state to produce the hidden state.

As shown in Formula 7, the sigmoid function (σ) is used as the activation function for the gates i_t , h_t , and o_t to ensure that their values stay within the range of 0 to 1 [65].

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{7}$$

The hyperbolic tangent function (\tanh) function, as shown in Formula 8, is used as an activation function for the cell state update to normalizes the input to a range between -1 and 1 , effectively centering the data and addressing vanishing gradient issues in neural networks. Its output is centered around 0 , and with a larger gradient near 0 , this helps the model to converge more quickly [65].

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{8}$$

We used the Rectified Linear Unit (ReLU) activation function for the fully connected layer, shown in Formula 9 to prevent the vanishing gradient problem, as it maintains a gradient of 1 for positive inputs, which helps in accelerating training and improving model performance [65].

$$\text{ReLU}(x) = \max(0, x) \tag{9}$$

The Softmax function (Formula 10) is used in the output layer for multi-class classification. It converts raw scores into probabilities for each of the 8 nutritional statuses, enabling the model to predict the likelihood of each category effectively.

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \tag{10}$$

Algorithmic explanation for LSTM-FC network architecture

The architecture of our LSTM-FC model, illustrated in Fig. S2, features two hidden layers with 256 and 128 neurons, and a fully connected layer with 128 neurons. The detailed algorithmic explanation is provided as follows:

Algorithm 1. LSTM-FC Neural Network Architecture

Phase 1: Input data preparation

1. Preprocess Input Data: Normalize and reshape the input data $X = \{X_1, X_2, \dots, X_n\}$ ensure compatibility with the model.

Phase 2: LSTM layer 1

2. Initialize LSTM₁: Define the first LSTM layer with 256 units using a tanh activation function.
3. Feed Input into LSTM₁: pass each feature X_i through LSTM₁, where the layer computes a sequence of 256-dimensional hidden states for each timestep.
4. Output of LSTM₁: Collect the output sequence, consisting of 256-dimensional hidden states.

Phase 3: LSTM layer 2

5. Initialize LSTM₂: Define the second LSTM layer with 128 units, utilizing a sigmoid activation function.
6. Feed Output of LSTM₁ into LSTM₂: Pass the 256-dimensional hidden state sequence from LSTM₁ to LSTM₂, generating a sequence of 128-dimensional hidden states.
7. Output of LSTM₂: Capture the output sequence of 128-dimensional hidden states.

Phase 4: Fully Connected (FC) layer

8. Initialize FC Layer: Define a fully connected layer with 128 neurons.
9. Feed Output of LSTM₂ into FC Layer: Pass the 128-dimensional output from LSTM₂ to the FC layer.
10. Output of FC Layer: Generate a 128-dimensional vector using the ReLU activation function.

Phase 5: Output layer

11. Initialize Output Layer: Configure the output layer with 8 units, corresponding to the target classes.
12. Feed Output of FC Layer into Output Layer: Pass the 128-dimensional output vector from the FC layer to the output layer.
13. Apply Softmax Activation: Compute class probabilities using the softmax activation function.

Phase 6: Training the network

14. Initialize Model Weights: Randomly initialize all model weights.
 15. Train the Model: Use backpropagation through time (BPTT) to minimize the loss function.
 16. Update Model Weights: Optimize the model weights using the Adam optimizer.
-

Model training strategy

During training, the dataset was split into training and testing sets, with the model trained over a maximum of 500 epochs. The model's dimensions were defined by batch size, time step, and feature count. The Adam optimizer was used with an initial learning rate of 0.001, and learning rate scheduling was applied to adjust the rate as training progresses [104]. To prevent overfitting, regularization techniques specifically dropout with rates ranging from 0.2 to 0.5 were employed. The training process is monitored using validation loss and accuracy, with early stopping implemented to halt training if there is no improvement after a specified number of epochs. Finally, the model was evaluated on the test set to assess its generalization ability, and hyperparameter tuning was performed to optimize its performance.

Hyperparameter tuning

Hyperparameters are key to model performance. Throughout the training phase, we explored a range of hyperparameters such as learning rate, number of layers, number of neurons, dropout rate, batch size, number of epochs, optimizer, and activation function. The learning rate controls weight update speed, batch size affects training efficiency, and the number of layers and units per layer shape the model's complexity. Dropout rate prevents overfitting, while epochs determine how often the dataset is fully processed. Proper tuning of these hyperparameters is essential for achieving optimal results. To determine the optimal number of hidden layers, we train models with varying configurations, starting with one layer and increasing complexity incrementally. We assess performance on a validation set and continue adding layers until gains are minimal. Similarly, for the number of neurons, we typically use powers of 2 (e.g., 64, 128, 256) for simplicity and ease of comparison.

Evaluating the model's ability to generalize

The generalization capability of the trained LSTM, which refers to its ability to accurately classify new and previously unseen data is crucial. To assess this, the classification error E^{class} is evaluated on the test set, which contains data that was not used during the training process. This allows for an unbiased evaluation of how well the LSTM model generalizes to new data. The classification error E^{class} is typically calculated by comparing the predicted labels with the true labels of the test set. A lower classification error indicates better generalization performance, as it suggests the model has learned patterns that can be applied to unseen instances. Furthermore, we employed eightfold cross-validation, where the entire dataset was divided into eight subsets (folds). In each iteration, one-fold was used as the testing set while the remaining seven folds were used for training the model. This process was repeated for all folds, ensuring that every data point was used for both training and testing, providing a robust estimate of the model's performance across different data splits.

Model performance metrics

In this paper, we used several performance metrics to evaluate the effectiveness of our model. These metrics include accuracy, sensitivity (recall), specificity (precision), F_1 score, and confusion matrix, each providing a distinct perspective on the model's

classification performance. The detailed definitions and explanations of these performance metrics, along with their respective formulas, are provided in the supplementary material. We evaluate the performance of the LSTM-FC model against several baseline methods: (I) the ARIMA, (II) RF, (III) RNN, (IV) the GRU, and (V) Bidirectional LSTM model, with detailed descriptions provided in the supplementary material.

Computational environment

In this study, we used Python’s TensorFlow and Keras libraries to leverage their powerful neural network capabilities for training and testing our LSTM-FC model [105] on a system equipped with an Intel I7-9300H CPU, NVIDIA GTX 1650 GPU, and 32 GB RAM.

Results

The nutritional status of children in Ethiopia, as detailed in Fig. 3, shows significant fluctuations over time. During infancy (<2 years), the proportion of children with a normal nutritional status peak, with 19.37% of females and 18.32% of males in this category. However, by early adolescence (12–14 years), this figure decreases to 15.12% for females and 16.63% for males. In contrast, as children grow older, severe undernutrition (USW), which includes underweight, stunted, and wasted conditions concurrently, increases. For males, this issue rises sharply during middle adolescence (14–18 years), affecting 47.59%. In contrast, females experience the highest rate of severe undernutrition earlier, with 52.34% affected during early adolescence (12–14 years). The detailed values can be found in the supplementary material, Table S4. Figure 4.

Comparison of metrics across different dataset splits

The dataset, collected across five time points was divided into various training and testing splits to evaluate model performance under different scenarios. Among the evaluated

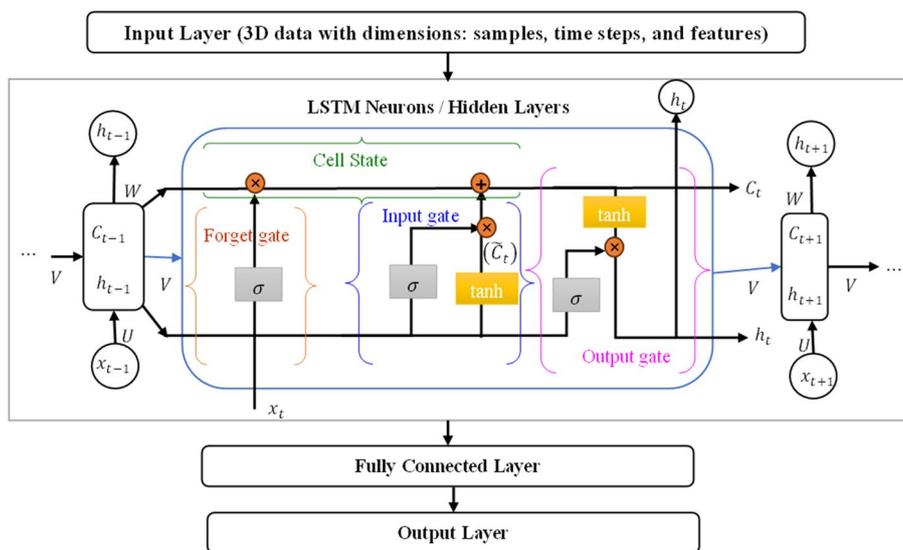


Fig. 3 Computational diagram of a LSTM-FC Cell Architecture [102, 106]

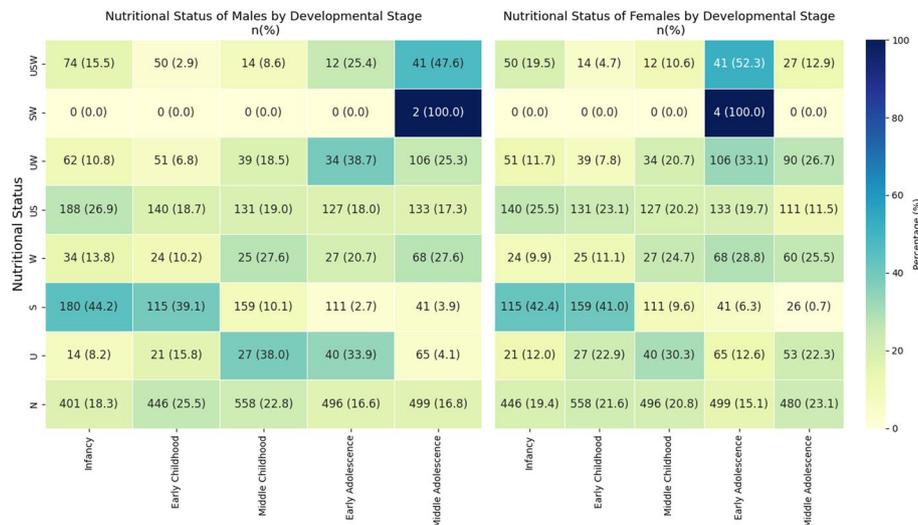


Fig. 4 Distribution of child nutritional status by their developmental stage and gender in Ethiopia (2002 – 2016). The values in parentheses give the percentage (%)

splits, training on the comprehensive dataset from 2002, 2006, 2009, and 2013 and testing on 2016 yielded the best performance, with high sensitivity (0.85), specificity (0.84), F1-score (0.84), AUC of 0.90 and an accuracy of 0.83 (Table 1).

Performance evaluation of LSTM-FC model against various baseline methods

In this section, we assess the predictive accuracy and long-term forecasting capability of the LSTM-FC model for evaluating child nutritional status. The model’s performance is compared across various developmental stages, including infancy, early childhood, middle childhood, early adolescence, and middle adolescence, against several baseline models (Fig. 5).

Our findings reveal that neural network-based approaches, particularly the LSTM-FC and Bi-LSTM models, excel in capturing temporal features and consistently outperform traditional baseline models such as RE, ARIMA, RNN, and GRU in terms of prediction accuracy and precision. For example, in the 2016 nutrition status forecasting task, the LSTM-FC and Bi-LSTM models achieve approximately 36% and 34% higher accuracy than the ARIMA model, with precision improvements of about 25% and 23%, respectively (Fig. 5). Additionally, these models exhibit substantial enhancements across key performance metrics, including accuracy, precision, recall, and F1-Score, when compared to the baseline methods.

The LSTM-FC model demonstrates superior performance across various evaluation metrics and prediction horizons, as illustrated in Fig. 5. It consistently achieves the highest metrics, regardless of the prediction horizon, and shows minimal fluctuations in results. This stability indicates that the model is robust and unaffected by horizon variations, making it suitable for both short-term and long-term predictions. Detailed performance values for each model are provided in Table S6 of the supplementary material for reference. Overall, the LSTM-FC model consistently outperformed ARIMA, RE, RNN, GRU, and Bi-LSTM-FC across all time periods, achieving the highest overall accuracy

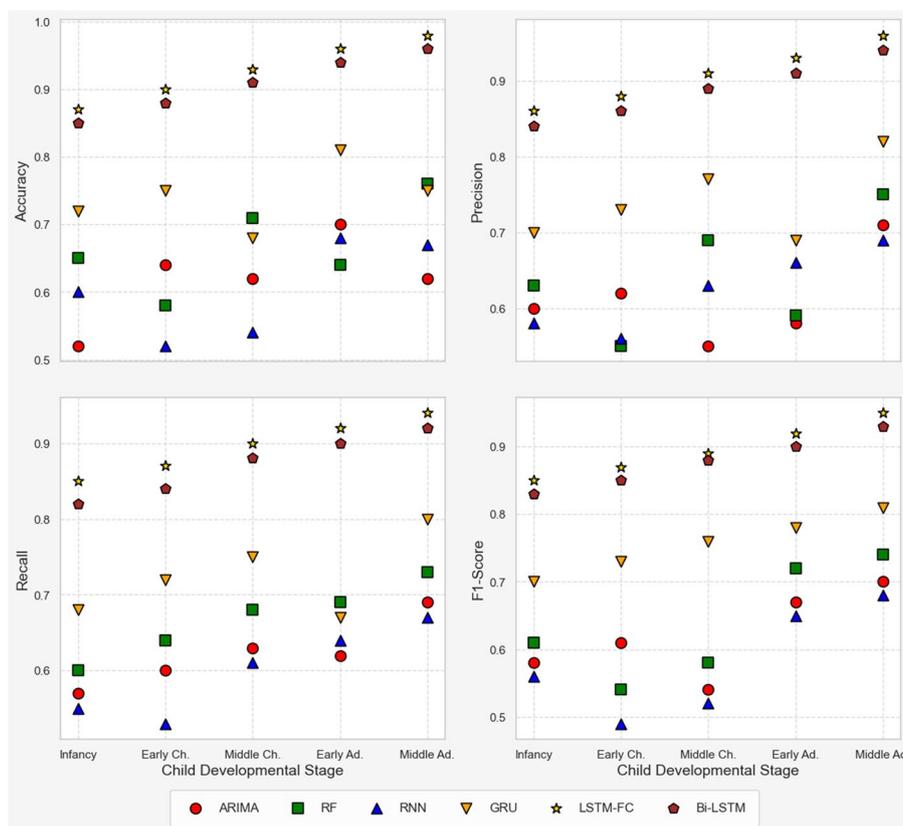


Fig. 5 The prediction results of the LSTM-FC model and other baseline prediction horizons on child nutritional dataset in Ethiopia

(93.4%), precision (91.4%), recall (89.8%), and F1-score (90.2%), demonstrating its superior capability in evaluating child nutritional status.

Classification of child nutritional status across gender: confusion matrix analysis

To further analyze the model’s performance, we present the confusion matrix for each nutritional status category across gender (Fig. 6). This comparison highlights the model’s performance in predicting nutritional status combinations, revealing potential differences in classification accuracy between male and female children. Each cell represents the count of instances predicted for a specific category. The diagonal values indicate correct classifications, while the off-diagonal values represent misclassifications between different nutritional status categories.

The confusion matrix in Fig. 6 visualizes the performance of the classification model, comparing true versus predicted nutritional status categories across gender. The model performs exceptionally well, with the normal category (N) and SW showing zero misclassifications, as indicated by the absence of off-diagonal values in that row for both males and females. The diagonal values for other categories are also significantly higher than the off-diagonal values for both genders, further showcasing the LSTM-FC model’s accuracy in classification. Overall, the model demonstrates strong predictive accuracy for most nutritional status categories across gender.

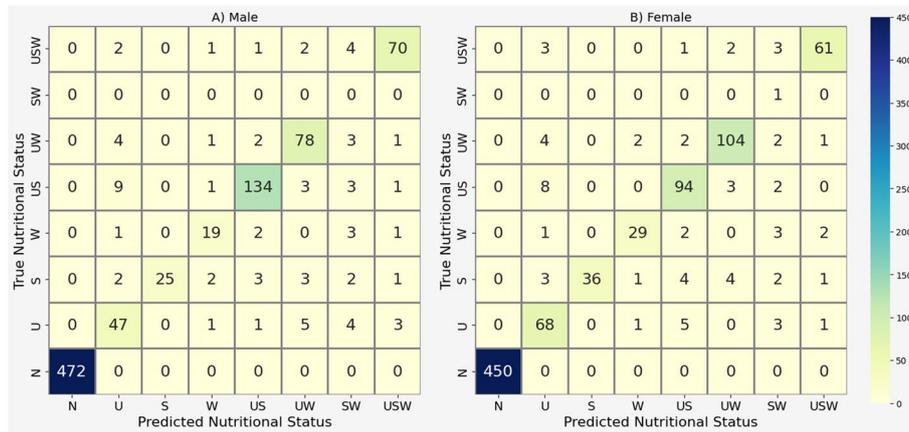


Fig. 6 Confusion matrix for child nutritional status classification model across gender

The model was trained to predict the nutritional status of children over 500 epochs, with a maximum of 4000 iterations, utilizing 8 iterations per epoch on a system equipped with a GPU. To gain a deeper understanding of the model’s performance, we analyze the prediction errors for each nutritional status category across gender, as shown in Fig. 7. Each point in the plot represents the prediction error for a specific category, where the error is the difference between the true and predicted nutritional status. The distribution of errors highlights the model’s performance, with certain categories showing tighter error ranges compared to others.

The N (normal) and SW (stunted and wasted concurrently) categories exhibit a mean error close to zero, indicating accurate predictions with minimal bias, consistent with the findings in Fig. 6. In contrast, most categories show negative errors in the range of 15 to -20, suggesting slight underestimation of nutritional status in those categories. With the exception of the U (underweight) and UW (underweight and wasted concurrently) categories, all other nutritional status categories exhibit negative prediction errors for both genders. Specifically, in females, the US (underweight and stunted simultaneously) category demonstrates a notably large prediction error, whereas in males, the S (stunted) category exhibits the largest prediction error.

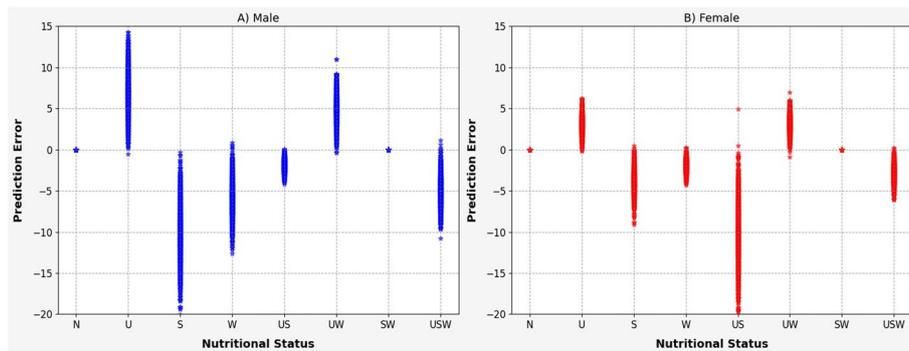


Fig. 7 Performance of nutritional status prediction across gender

Prevalence prediction percentage

The plot presents the nutritional status of children across different age groups, with the percentage of children affected by various conditions. For instance, the normal (N) category shows a gradual decline in the percentage of children with healthy nutritional status from 46.2% at age 1 to 32.7% at age 25. The underweight (U) category fluctuates slightly, peaking at 5.9% at age 8 and declining to 2% at age 1 and 15% at age 25. The stunted (S) condition peaks at 50.8% for children between the ages of 8 and 15, gradually declining to 3.8% by age 25. The wasted (W) condition starts at 3.2%, peaks at 7.2% at age 15, and stabilizes around 6.9% at age 25. In the second subplot, the underweight and stunted (US) category fluctuates, with a peak of 18.6% at age 1 and a decline after the early years. The underweight and wasted (UW) category starts at 6.1% at age 1, increases to 19.4% at age 15, and then decreases. The stunted and wasted (SW) category remains low, with percentages ranging from 0 to 0.7%. The underweight, stunted, and wasted (USW) category starts at 7.4% at age 1, peaks at 14.4% at age 12, and ends at 9.2% by age 25 (Fig. 8).

Evaluation of model’s generalizability

The progression of E^{class} , assessed on the training dataset, is shown in Fig. 9(a). In runs incorporating both classification and prediction task, the error decreases significantly more rapidly, resulting in a much smaller classification error after training. Figure 9(b) presents the success rate of finding a solution with $E^{class} \leq 5\%$, as a function of the maximum number of epochs. Success rate is measured on both the training and test datasets. When incorporating the additional source of information, 96% of the runs achieve an E^{class} on the training set reaches the 5%, and 94% of the runs meet this threshold on the test set. These results indicate a substantial improvement in model generalizability, as the rate is almost higher than when the training process relies solely on the classification task.

To further evaluate the generalizability of the model, eightfold cross-validation was performed, and the evaluation results for each fold are reported in Table 2. This approach ensures that the model’s performance is assessed on diverse subsets of the data, providing a comprehensive understanding of its generalization ability. The results

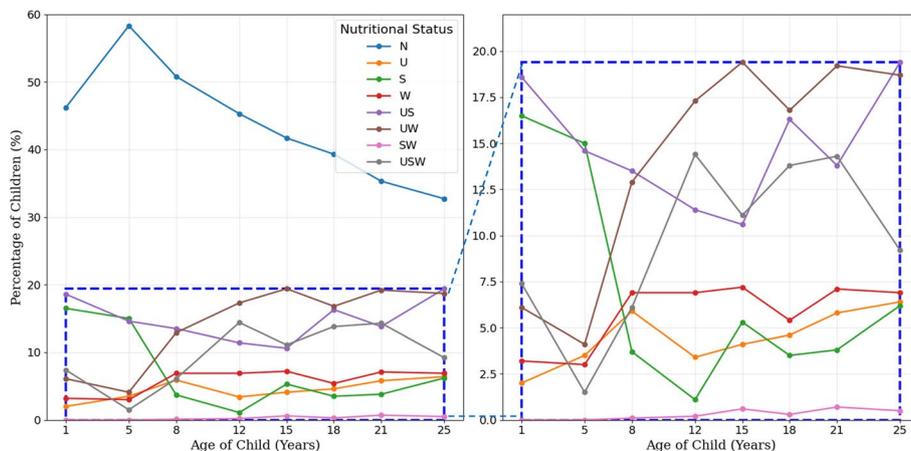


Fig. 8 Prevalence prediction of Ethiopian children nutritional status from age 1 to 25

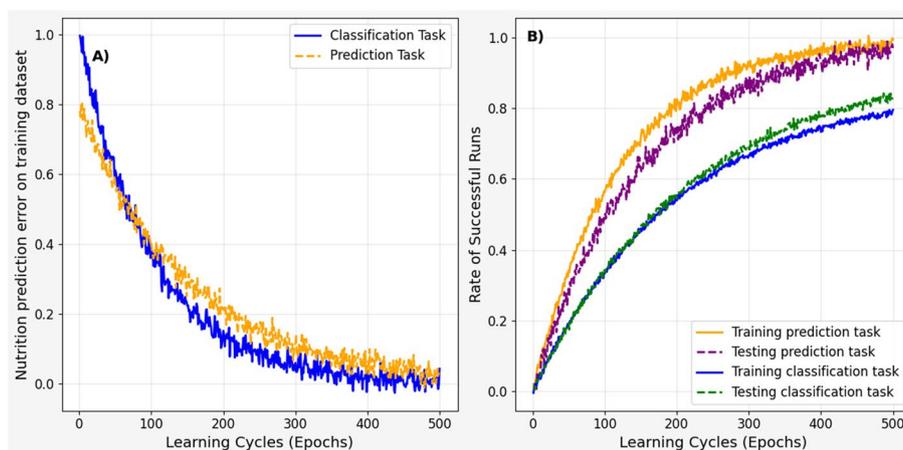


Fig. 9 Generalizability of LSTM-FC Model

Table 2 Generalization performance scores of the LSTM-FC model

Metrics	Testing	Generalization Evaluation				
		Fold 4	Fold 5*	Fold 6	Fold 7	Fold 8
Precision	0.68319	0.580241	0.659877	0.612346	0.593226	0.602658
Recall	0.92841	0.978764	0.993714	0.975116	0.948571	0.984286
Accuracy	0.65295	0.580241	0.642169	0.606024	0.580000	0.598072
F1-Score	0.77592	0.738147	0.798413	0.769355	0.752720	0.769421

Note: Fold 5 is highlighted (*) as the best-performing fold due to its superior Precision, Recall, Accuracy, and F1-Score compared to other folds. The Testing column represents the evaluation of the model's performance on an unseen test dataset

show varying performance across the folds, with Fold 5 consistently achieving the highest scores in Precision, Recall, Accuracy, and F1-Score, indicating its superior generalization capability compared to the other folds. These findings are further supported by the training results and visualized in Fig. 9, demonstrating the model's effectiveness in both classification and prediction tasks.

Discussion

Previous research indicates that Ethiopia ranks among the countries with the highest number of undernourished children globally and in Sub-Saharan Africa [107–112]. Although there has been a decline in under-five undernutrition in Ethiopia, it has not met Millennium Development Goals expectations [112]. Undernutrition in the first 1000 days (from birth to 23 months) often persists until age five and beyond, leading to long-term physical, psychological, cultural, and socioeconomic impacts [113]. Furthermore, the prevalence of undernutrition into later developmental stages is under-researched globally, including in Ethiopia [114, 115].

In comparing the performance of LSTM-FC across various predictive horizons, our study demonstrates the clear superiority of the LSTM-FC model over other baseline models, including ARIMA, RF, RNN, GRU, and Bi-LSTM-FC, at all-time points. The LSTM-FC consistently achieves higher metrics, emphasizing its effectiveness in capturing temporal dependencies and complex sequential patterns in longitudinal data. These

findings align with existing literature [116, 117], which highlights the strength of LSTM models in retaining long-term dependencies. While models like GRU and Bi-LSTM-FC also perform well, their metrics fall slightly short of LSTM-FC, particularly in precision and recall, indicating a less balanced trade-off between false positives and false negatives.

In a study conducted by Sima Siami-Namini et al. [118], Bi-LSTM demonstrated superior predictive performance compared to ARIMA and LSTM models; however, in our study, LSTM-FC outperformed Bi-LSTM with slightly better prediction, highlighting its enhanced ability to model temporal dependencies in longitudinal data. Similarly, in the study by Pirani et al. [119], a comparative analysis of ARIMA, GRU, LSTM, and Bi-LSTM was conducted on financial time series forecasting. The authors demonstrated that LSTM-based models, particularly Bi-LSTM, outperformed traditional methods like ARIMA. While Bi-LSTM provides enhanced training with its bidirectional structure, the study highlights the importance of choosing the right algorithm for minimizing errors and achieving superior predictions. In contrast, our study finds that LSTM-FC slightly outperforms Bi-LSTM, further emphasizing the effectiveness of LSTM-based architectures in time-series prediction tasks.

Optimizing the number of hidden layers and neurons is crucial for enhancing our LSTM-FC model's performance and efficiency. More layers help capture complex patterns, and more neurons allow for processing extensive information [29]. The right balance prevents overfitting, reduces computational demands, and shortens training times [120]. Through extensive testing, we identified the optimal layer and neuron configuration for our dataset. As a result, the LSTM-FC model's architecture, featuring two hidden layers followed by a fully connected layer, strikes an optimal balance between complexity and computational efficiency. This design, consistent with findings that two to three hidden layers are typically sufficient for most applications [26, 65, 121–123].

The research conducted by Amin and Novitasari [124] used an LSTM model to identify stunting in hospital data based on gender, age, and height. This approach leverages the LSTM's ability to handle sequential data, aiding in the early detection and diagnosis of stunting. However, in our study, combining LSTM with fully connected (FC) layers outperformed the standard LSTM and bidirectional models, yielding better results for predicting child nutritional status. Addressing overfitting challenges, our model uses a single FC layer with a dropout ratio of 0.3, differing from Zhao et al.'s approach of using two FC layers with a dropout of 0.5 and 0.3 [65].

In conclusion, the use of LSTM-FC in healthcare is still relatively limited. Notably, Zhao et al. [65] employed a deep neural network combining LSTM and FC layers to predict bronchopneumonia in pediatric patients. Additionally, LSTMs have been effectively used in other applications, including the identification of surgical site infections (SSIs) from electronic health records (EHRs) [125]. Similarly, our research demonstrates the LSTM-FC model's superior performance in handling long-term dependencies in time-series data, highlighting its potential in the dynamic field of nutritional epidemiology. However, the implementation of the LSTM-FC model does require significant computational resources, which may pose challenges in low-resource settings.

This paper was aimed to predict the prevalence of nutritional status of children beyond age 15 using the LSTM-FC model. It analyzes data from the Young Lives study in Ethiopia, which tracked children's nutritional status from age 1 in 2002 to age 15 in 2016 through five survey rounds, with data collected approximately every 3.5 years. This study serves as a benchmark, demonstrating the effective implementation of LSTM-FC algorithms to predict nutritional status as children mature.

Conclusion

Research on the early prediction of children's nutritional status is crucial, as undernutrition remains a critical global health issue, particularly affecting developing countries such as Ethiopia. To the best of our knowledge, this study is the first to investigate the use of an LSTM-FC neural network model for classifying and predicting nutritional status of Ethiopian children, showcasing its superior performance over baseline models. The LSTM-FC model achieves high accuracy in classification, with most predictions correct and only a few misclassifications observed in both male and female categories. It performs particularly well in identifying normal and stunted-wasted (SW) nutritional statuses. Furthermore, the prediction error is also minimal, with a slight underestimation observed, specifically for the stunted category in males and the US category in females. Despite this, the overall classification accuracy remains strong across both genders. Likewise, the evaluation of the model's generalizability highlights its robust performance across various nutritional categories, with Fold 5 yielding the best results, while minimal prediction errors. Finally, our findings reveal a concerning decline in the healthy nutritional status of children, with a notable increase in the prevalence of multiple undernutrition conditions by 2027. These results underscore the urgent need for targeted public health interventions and policies aimed at combating undernutrition among Ethiopian children. Given Ethiopia's challenges with healthcare facilities, food insecurity, and displacement due to political instability and natural disasters, it is imperative for the government, private sectors, and policymakers to take decisive action.

Strengths, limitations, and future work

We compare the performance of the LSTM-FC model with existing baseline methods, demonstrating its superiority in terms of classification accuracy, generalizability, and minimal prediction error across various nutritional categories. Additionally, the study allocates sufficient time for extensive hyperparameter tuning, ensuring the model achieves optimal performance with minimal error. Similarly, the paper underwent rigorous data preprocessing steps to ensure the quality and reliability of the dataset. Finally, the paper considers the concurrent nutritional conditions of children, accounting for multiple undernutrition states simultaneously, which provides a more comprehensive understanding of their nutritional status.

One limitation of this study is the lack of comparison with hybrid models such as LSTM-RE, LSTM-ARIMA, and LSTM-CNN. These models are highlighted in the literature for their strong predictive capabilities. Additionally, advanced spatio-temporal models like T-GCN (Temporal Graph Convolutional Network) are known for their

ability to capture both spatial and temporal dependencies. Future research should explore these models to assess their performance against LSTM-FC and other baseline models. This could provide deeper insights into their applicability for time-series prediction and spatio-temporal analysis tasks. Furthermore, future research with the LSTM-FC model aims to first ensure access to a system equipped with a high-performance GPU to facilitate efficient hyperparameter tuning, enabling improved optimization and performance evaluation.

Abbreviations

ARIMA	Auto-Regressive Integrated Moving Average
Bi-LSTM	Bidirectional LSTM
EAP	Emergency Aid Program
GRU	Gated Recurrent Unit
HEP	Health Extension Program
LSTM	Long Short-Term Memory
PSNP	Productive Safety Net Program
RNN	Recurrent Neural Network
RF	Random Forest
SHAP	Shapley Additive Explanations
SNNP	Southern Nations, Nationalities, and Peoples
SSA	Sub-Saharan Africa
SMOTE	Synthetic Minority Oversampling Technique
WHO	World Health Organization
YLCS	Young Lives Cohort Survey

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13040-025-00425-0>.

Additional file 1. The supplementary material for this study provides additional details that support the findings in the main manuscript. This includes comprehensive tables and figures on the distribution of variables used in the analysis. Further, detailed bivariate analysis is presented to explore the relationships between key variables. Additionally, the supplementary material also includes performance metrics for the LSTM-FC and baseline models across time points, as well as classification and prediction accuracy for nutritional status, highlighting prediction errors and model performance for different categories.

Acknowledgements

The authors express their sincere gratitude to the reviewers for their insightful comments and suggestions, which greatly improved this manuscript. We also thank the Young Lives Study teams and the UK Data Service for providing access to the data files used in this research.

Artificial Intelligence (AI)-assisted technologies

We confirm that this manuscript was produced without the utilization of AI-assisted technologies, including Large Language Models (LLMs), chatbots, image creators, or any other AI-based tools.

Authors' contributions

GBB: Contributed to the conceptualization, conducted the formal analysis, shaped the methodology, performed the data analysis, engaged in the writing of the report, and authored the original draft. TZ: Contributed to conceptualization, guided methodology, reviewed and edited the manuscript, contributed to data analysis, and supervised the entire process from start to finish, ultimately approving the final manuscript. HMF: Provided valuable contributions to the methodology by refining its approach, offering expertise in software utilization, aiding in report writing, conducting comprehensive data analysis, and providing insightful suggestions to elevate the overall quality of the research.

Funding

This research has no specific grant from any funding agency, commercial, or not-for-profit sectors to report.

Data availability

The dataset used in this study was obtained from the Young Lives Study. Access to the data can be obtained either by completing the form available at <https://www.younglives.org.uk/use-our-data-form> selecting the dataset "Young Lives: Rounds 1-5 constructed files, 2002-2016" (used in this study), or by creating a user account through the <https://ukdataservice.ac.uk/> subject to their terms and conditions. Additionally, the survey questionnaires for each round (Rounds 1-5) are available through the following link: <https://www.younglives.org.uk/round-1-questionnaires>. By adjusting the round number in the URL or navigating through the menu on the Young Lives website, users can access the questionnaires for each respective round.

Declarations

Ethics approval and consent to participate

Ethics approval was not required for this study as it involved secondary data analysis of publicly available datasets. No new human participants were recruited, and no direct human or animal subjects were involved. The data were anonymized and publicly accessible, ensuring no ethical approval was necessary.

Consent for publication

As this study utilizes publicly available datasets, individual consent for publication was not required. The data used in this research were anonymized, ensuring that no personal identifiers were included, in accordance with ethical research practices.

Competing interests

The authors declare no competing interests.

Received: 15 September 2024 Accepted: 17 January 2025

Published online: 30 January 2025

References

1. Bourke CD, Jones KD, Prendergast AJ. Current understanding of innate immune cell dysfunction in childhood undernutrition. *Front Immunol.* 2019;10:1728.
2. Morales F, et al. Effects of malnutrition on the immune system and infection and the role of nutritional strategies regarding improvements in children's health status: A literature review. *Nutrients.* 2023;16(1):1.
3. Calder PC, Jackson AA. Undernutrition, infection and immune function. *Nutr Res Rev.* 2000;13(1):3–29.
4. Khan S, Zaheer S, Safdar NF. Determinants of stunting, underweight and wasting among children < 5 years of age: evidence from 2012–2013 Pakistan demographic and health survey. *BMC Public Health.* 2019;19:1–15.
5. Gupta AK, Borkotoky K. Exploring the multidimensional nature of anthropometric indicators for under-five children in India. *Indian J Public Health.* 2016;60(1):68–72.
6. Kassie GW, Workie DL. Exploring the association of anthropometric indicators for under-five children in Ethiopia. *BMC Public Health.* 2019;19:1–6.
7. Amadu I, et al. Risk factors associated with the coexistence of stunting, underweight, and wasting in children under 5 from 31 sub-Saharan African countries. *BMJ Open.* 2021;11(12): e052267.
8. Emerson E, Savage A, Llewellyn G. Prevalence of underweight, wasting and stunting among young children with a significant cognitive delay in 47 low-income and middle-income countries. *J Intellect Disabil Res.* 2020;64(2):93–102.
9. Bisai S, Ghosh T, Bose K. Prevalence of underweight, stunting and wasting among urban poor children aged 1–5 years of West Bengal. *India Int J Curr Res.* 2010;6:39–44.
10. Ntenda PAM. Association of low birth weight with undernutrition in preschool-aged children in Malawi. *Nutr J.* 2019;18:1–15.
11. Toma TM, et al. Factors associated with wasting and stunting among children aged 06–59 months in South Ari District, Southern Ethiopia: a community-based cross-sectional study. *BMC nutrition.* 2023;9(1):34.
12. United Nations. The State of the World's Children 2019: Children, Food and Nutrition - Growing Well in a Changing World. State of the World's Children. 2019. <https://doi.org/10.18356/a4da4811-en>.
13. Thurstans S, et al. The relationship between wasting and stunting in young children: A systematic review. *Matern Child Nutr.* 2022;18(1): e13246.
14. De Onis M, Blössner M, Borghi E. Prevalence and trends of stunting among pre-school children, 1990–2020. *Public Health Nutr.* 2012;15(1):142–8.
15. UN. The Millennium Development Goals Report 2012. Millennium Development Goals Report. 2012. <https://doi.org/10.18356/32f1e244-en>.
16. Organization, W.H., The state of food security and nutrition in the world 2020: transforming food systems for affordable healthy diets. Vol. 2020. 2020: Food & Agriculture Org.
17. Sachs J, Lafortune G, Fuller G. Sustainable Development Report 2019. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN); 2019. Available at: https://s3.amazonaws.com/sustainabledevelopment.report/2019/2019_sustainable_development_report.pdf.
18. Salam RA, Das JK, Bhutta ZA. Current issues and priorities in childhood nutrition, growth, and infections. *J Nutr.* 2015;145(5):1116S–1122S.
19. Che Z, et al. Recurrent neural networks for multivariate time series with missing values. *Sci Rep.* 2018;8(1):6085.
20. Bharadiya JP. Exploring the use of recurrent neural networks for time series forecasting. *International Journal of Innovative Science and Research Technology.* 2023;8(5):2023–7.
21. Hu B. Research on Natural Language Processing Problems Based on LSTM Algorithm. Proceedings of the 3rd Asia-Pacific Conference on Image Processing, Electronics and Computers. 2022. p. 259–63.
22. Oruh J, Viriri S, Adegun A. Long short-term memory recurrent neural network for automatic speech recognition. *IEEE Access.* 2022;10:30069–79.
23. Espinosa R, et al. A time series forecasting based multi-criteria methodology for air quality prediction. *Appl Soft Comput.* 2021;113: 107850.

24. Allam F, Nossai Z, Gomma H, Ibrahim I, Abdelsalam M. A recurrent neural network approach for predicting glucose concentration in type-1 diabetic patients. In International Conference on Engineering Applications of Neural Networks, Berlin, Heidelberg: Springer Berlin Heidelberg; 2011. p. 254-259.
25. Villegas M, et al. Predicting the evolution of COVID-19 mortality risk: A Recurrent Neural Network approach. *Computer Methods and Programs in Biomedicine Update*. 2023;3: 100089.
26. Shahriar MM, Iqbal MS, Mitra S, Das AK. A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month's Older Children in Bangladesh. In 2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT). IEEE; 2019; p. 145-149.
27. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. 1997;9(8):1735–80.
28. Hochreiter S. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *Internat J Uncertain Fuzziness Knowledge-Based Systems*. 1998;6(02):107–16.
29. Bengio Y, Simard P, Frasconi P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans Neural Networks*. 1994;5(2):157–66.
30. Asfaw M, et al. Prevalence of undernutrition and associated factors among children aged between six to fifty nine months in Bule Hora district. *South Ethiopia BMC Public Health*. 2015;15:1–9.
31. Hovhannisyan L, Demirchyan A, Petrosyan V. Estimated prevalence and predictors of undernutrition among children aged 5–17 months in Yerevan. *Armenia Public health nutrition*. 2014;17(5):1046–53.
32. Medhin G, et al. Prevalence and predictors of undernutrition among infants aged six and twelve months in Butajira, Ethiopia: the P-MaMiE Birth Cohort. *BMC Public Health*. 2010;10:1–15.
33. Ricci, C., et al. Determinants of undernutrition prevalence in children aged 0–59 months in sub-Saharan Africa between 2000 and 2015. A report from the World Bank database. *Public health nutrition*, 2019. 22(9):1597–1605.
34. Ijaiya MA, Anjorin S, Uthman OA. Quantifying the increased risk of illness in malnourished children: a global meta-analysis and propensity score matching approach. *Global Health Research and Policy*. 2024;9(1):29.
35. Das S, Baffour B, Richardson A. Prevalence of child undernutrition measures and their spatio-demographic inequalities in Bangladesh: an application of multilevel Bayesian modelling. *BMC Public Health*. 2022;22(1):1008.
36. Muhe A, et al. Predictors of stunting among children age 6–59 months in Ethiopia using Bayesian multi-level analysis. *Sci Rep*. 2021;11(1):3759.
37. Takele BA, Gezie LD, Alamneh TS. Pooled prevalence of stunting and associated factors among children aged 6–59 months in Sub-Saharan Africa countries: A Bayesian multilevel approach. *PLoS ONE*. 2022;17(10): e0275889.
38. Egata G, Berhane Y, Worku A. Seasonal variation in the prevalence of acute undernutrition among children under five years of age in east rural Ethiopia: a longitudinal study. *BMC Public Health*. 2013;13:1–8.
39. Wake SK, et al. Longitudinal trends and determinants of stunting among children aged 1–15 years. *Archives of Public Health*. 2023;81(1):60.
40. Owoeye SM, Oseni BM, Gayawan E. Estimating lifetime malnourished period and its statistics based on the concept of Markov chain with reward. *Heliyon*. 2020;6(5):e04073.
41. Begashaw GB, Zewotir T, Fenta HM. Multistate Markov chain modeling for child undernutrition transitions in Ethiopia: a longitudinal data analysis, 2002–2016. *BMC Med Res Methodol*. 2024;24(1):283.
42. How, M.-L. and Y.J. Chan, Artificial intelligence-enabled predictive insights for ameliorating global malnutrition: a human-centric ai-thinking approach. *AI*, 2020. 1(1): 4.
43. Begashaw, G.B., T. Zewotir, and H. Mekonnen, Dynamic Bayesian Network Modeling for Longitudinal Data on Child Undernutrition in Ethiopia (2002–2016). *Frontiers in Public Health*. 12: p. 1399094.
44. Bitew FH, Sparks CS, Nyarko SH. Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public Health Nutr*. 2022;25(2):269–80.
45. Fenta HM, Zewotir T, Muluneh EK. A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Med Inform Decis Mak*. 2021;21:1–12.
46. Anku EK, Duah HO. Predicting and identifying factors associated with undernutrition among children under five years in Ghana using machine learning algorithms. *PLoS ONE*. 2024;19(2): e0296625.
47. Jain S, Khanam T, Abedi AJ, Khan AA. Efficient Machine Learning for Malnutrition Prediction among under-five children in India. In 2022 IEEE Delhi Section Conference (DELCON). IEEE; 2022. p. 1-10.
48. Mambang M, Marleny FD, Zulfadhilah M. Prediction of linear model on stunting prevalence with machine learning approach. *Bulletin of Electrical Engineering and Informatics*. 2023;12(1):483–92.
49. Kar S, Pratihar S, Nayak S, Bal S, Gururaj HL, Ravikumar V. Prediction of child malnutrition using machine learning. In 2021 10th International Conference on Internet of Everything, Microwave Engineering, Communication and Networks (IEMECN). IEEE; 2021. p. 01-04.
50. Talukder A, Ahammed B. Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*. 2020;78: 110861.
51. Singer P, et al. Detecting memory and structure in human navigation patterns using Markov chain models of varying order. *PLoS ONE*. 2014;9(7): e102070.
52. Greenan KM, Plank JS, Wylie JJ. Mean Time to Meaningless: {MTTDL}, Markov Models, and Storage System Reliability. In 2nd Workshop on Hot Topics in Storage and File Systems (HotStorage 10). 2010.
53. De Blasi RA, Campagna G, Finazzi S. A dynamic Bayesian network model for predicting organ failure associations without predefining outcomes. *PLoS ONE*. 2021;16(4): e0250787.
54. Wu PP-Y, et al. Dynamic Bayesian network inferencing for non-homogeneous complex systems. *J R Stat Soc: Ser C: Appl Stat*. 2018;67(2):417–34.
55. Hochreiter, S. and J. Schmidhuber, LSTM can solve hard long time lag problems. *Adv neural inform process syst*. 1996. 9.

56. DiPietro, R. and G.D. Hager. Automated surgical activity recognition with one labeled sequence. in Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part V 22. 2019. Springer.
57. Navares R, Aznarte JL. Predicting air quality with deep learning LSTM: Towards comprehensive models. *Eco Inform.* 2020;55: 101019.
58. Verma, H. and S. Kumar. An accurate missing data prediction method using LSTM based deep learning for health care. in Proceedings of the 20th international conference on distributed computing and networking. 2019.
59. Al-Selwi SM, et al. LSTM inefficiency in long-term dependencies regression problems. *Journal of Advanced Research in Applied Sciences and Engineering Technology.* 2023;30(3):16–31.
60. Lechner, M. and R. Hasani, Learning long-term dependencies in irregularly-sampled time series. arXiv preprint [arXiv:2006.04418](https://arxiv.org/abs/2006.04418), 2020.
61. Ravi D, et al. Deep learning for health informatics. *IEEE J Biomed Health Inform.* 2016;21(1):4–21.
62. Lipton, Z., Learning to diagnose with LSTM recurrent neural networks. arXiv preprint [arXiv:1511.03677](https://arxiv.org/abs/1511.03677), 2015.
63. Greff K, et al. LSTM: A search space odyssey. *IEEE transactions on neural networks and learning systems.* 2016;28(10):2222–32.
64. Athanasiou M, et al. Long short-term memory–based prediction of the spread of influenza-like illness leveraging surveillance, weather, and twitter data: Model development and validation. *J Med Internet Res.* 2023;25: e42519.
65. Zhao D, et al. A long short-term memory-fully connected (LSTM-FC) neural network for predicting the incidence of bronchopneumonia in children. *Environ Sci Pollut Res.* 2021;28(40):56892–905.
66. Yang S, Chen H-C, Wu C-H, Wu M-N, Yang C-H. Forecasting of Dementia Using the LSTM Neural Network in Taiwan. *Mathematics.* 2021;9(5):488. <https://doi.org/10.3390/math9050488>.
67. Dai, Z., S. Liu, and C. Liu, Predict the prevalence and incidence of Parkinson’s disease using fractal interpolation-LSTM model. *Chaos: An Interdisciplinary J Nonlinear Sci.* 2024. 34(5).
68. Van VTS, et al. Predicting undernutrition among elementary schoolchildren in the Philippines using machine learning algorithms. *Nutrition.* 2022;96: 111571.
69. Shen H, Zhao H, Jiang Y. Machine learning algorithms for predicting stunting among under-five children in Papua New Guinea. *Children.* 2023;10(10):1638.
70. Chilyabanyama ON, et al. Performance of machine learning classifiers in classifying stunting among under-five children in Zambia. *Children.* 2022;9(7):1082.
71. Khan JR, Tomal JH, Raheem E. Model and variable selection using machine learning methods with applications to childhood stunting in Bangladesh. *Inform Health Soc Care.* 2021;46(4):425–42.
72. Ndagijimana S, et al. Prediction of stunting among under-5 children in Rwanda using machine learning techniques. *J Prev Med Public Health.* 2023;56(1):41.
73. Rahman SJ, et al. Investigate the risk factors of stunting, wasting, and underweight among under-five Bangladeshi children and its prediction based on machine learning approach. *PLoS ONE.* 2021;16(6): e0253172.
74. Kawo MA, et al. Predicting undernutrition risk factors using machine learning techniques in Nigerian under five children. *International Journal of Computer Science and Mobile Computing.* 2024;13(7):56-70.
75. Biradar, V.G. and K.K. Naik. Classification of Children as Malnourished and Healthy Category Using ResNet18 Model. in 2023 International Conference on Ambient Intelligence, Knowledge Informatics and Industrial Electronics (AIKIE). 2023. IEEE.
76. Khare S, et al. Investigation of nutritional status of children based on machine learning techniques using Indian demographic and health survey data. *Procedia computer science.* 2017;115:338–49.
77. Vasu SR, Khare S, Gupta D, Jyotishi A. Features Explaining Malnutrition in India: A Machine Learning Approach to Demographic and Health Survey Data. In *Advanced Computing.* Singapore: Springer; 2021. p. 87–99. https://doi.org/10.1007/978-981-16-0401-0_7.
78. Von Grafenstein, L., Mapping Real-Time Child Malnutrition in India Using Machine Learning. 2023.
79. Kebede Kassaw A, et al. The application of machine learning approaches to determine the predictors of anemia among under five children in Ethiopia. *Sci Rep.* 2023;13(1):22919.
80. Sasaki R, Nakao Y, Mawatari F, Nishihara T, Haraguchi M, Fukushima M, Nakao K. Investigation of a chest radiograph-based deep learning model to identify an imaging biomarker for malnutrition in older adults. *Clinical Nutrition Open Science.* 2024;58:240-51.
81. Lakshminarayanan AR, Pavani B, Rajeswari V, Parthasarathy S, Khan AAA, Sathick KJ. Malnutrition detection using convolutional neural network. In 2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII). IEEE; 2021. p. 1-5.
82. Biradar VG, Hanumanthappa H, Pareek PK, Sravanthi T, Aishwarya K, Aneeth K. Performance Analysis of Deep Learning Models for Detection of Malnutrition in Children. In 2024 International Conference on Data Science and Network Security (ICDSNS). IEEE; 2024. p. 1-8.
83. Biradar, V.G. and K.K. Naik. Detection of Malnutrition in Children Using Deep Learning Model. in International Conference on Smart Computing and Communication. 2024. Springer.
84. Ankalaki S, et al. A Deep Learning Approach for Malnutrition Detection. *Int J Online Biomed Eng.* 2024;20(6):116–38.
85. Jin BT, et al. Predicting malnutrition from longitudinal patient trajectories with deep learning. *PLoS ONE.* 2022;17(7): e0271487.
86. Boyden, J., Young lives: An international study of childhood poverty: Round 2 2006 (Study# SN 6852); Round 3 2009 (Study# SN 6853). 2012.
87. Lives Y. Equating Cognitive Scores across Rounds and Cohorts for Young Lives in Ethiopia. India: Peru and Vietnam; 2020.

88. Argaw BA. Regional inequality of economic outcomes and opportunities in Ethiopia: A tale of two periods. WIDER Working Paper. UNU-WIDER. 2017. <https://doi.org/10.35188/unu-wider/2017/342-4>.
89. Oxford, A Guide to Young Lives Research. , Y. Lives, Editor. 2017.
90. Porter C, Favara M, Woldehanna T. Smarter through Social Protection? Evaluating the Impact of Ethiopia's Safety-Net on Child Cognitive Abilities. IZA Working Papers; No. 10972. Institute of Labor Economics. 2017. <http://ftp.iza.org/dp10972.pdf>.
91. Wang, H., et al., Ethiopia health extension program: an institutionalized community approach for universal health coverage. 2016: World Bank Publications.
92. Fenta HM, Zewotir T, Muluneh EK. Disparities in childhood composite index of anthropometric failure prevalence and determinants across Ethiopian administrative zones. *PLoS ONE*. 2021;16(9): e0256726.
93. Yirga AA, et al. Factors affecting child malnutrition in Ethiopia. *Afr Health Sci*. 2019;19(2):1897–909.
94. Headey D. An analysis of trends and determinants of child undernutrition in Ethiopia, 2000–2011. *Gates Open Res*. 2019;3(983):983.
95. Ersino G, et al. Gender and household structure factors associated with maternal and child undernutrition in rural communities in Ethiopia. *PLoS ONE*. 2018;13(10): e0203914.
96. Sahiledengle B, et al. Determinants of undernutrition among young children in Ethiopia. *Sci Rep*. 2022;12(1):20945.
97. Degarege D, Degarege A, Animut A. Undernutrition and associated risk factors among school age children in Addis Ababa. *Ethiopia BMC public health*. 2015;15:1–9.
98. Fenta HM, et al. Factors of acute respiratory infection among under-five children across sub-Saharan African countries using machine learning approaches. *Sci Rep*. 2024;14(1):15801.
99. Sendra-Arranz R, Gutiérrez A. A long short-term memory artificial neural network to predict daily HVAC consumption in buildings. *Energy and Buildings*. 2020;216: 109952.
100. Gu X, et al. The sliding window and SHAP theory—an improved system with a long short-term memory network model for state of charge prediction in electric vehicle application. *Energies*. 2021;14(12):3692.
101. Rodrigues, E.M., Y. Baghoussi, and J. Mendes-Moreira, KDBI special issue: Explainability feature selection framework application for LSTM multivariate time-series forecast self optimization. *Expert Systems: e13674*.
102. Kim S. Predicting Disease Progression Using Deep Recurrent Neural Networks and Longitudinal Electronic Health Record Data. 2020. p. 117203.
103. Carrillo-Moreno J, et al. Long short-term memory neural network for glucose prediction. *Neural Comput Appl*. 2021;33:4191–203.
104. Kingma, D.P. and J. Ba, Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
105. Lee T, et al. Tensorflow and keras programming for deep learning. *Deep Learn Hydrometeorol Environ Sci*. 2021;99:151–62.
106. Vennerød, C.B., A. Kjærran, and E.S. Bugge, Long short-term memory RNN. *arXiv preprint arXiv:2105.06756*, 2021.
107. Fenta HM, et al. Determinants of stunting among under-five years children in Ethiopia from the 2016 Ethiopia demographic and Health Survey: Application of ordinal logistic regression model using complex sampling designs. *Clinical Epidemiology and Global Health*. 2020;8(2):404–13.
108. Takele K, Zewotir T, Ndanguza D. Understanding correlates of child stunting in Ethiopia using generalized linear mixed models. *BMC Public Health*. 2019;19:1–8.
109. Gebre A, et al. Prevalence of malnutrition and associated factors among Under-five children in pastoral communities of Afar Regional State, Northeast Ethiopia: a community-based cross-sectional study. *Journal of nutrition and metabolism*. 2019;2019(1):9187609.
110. Kassie GW, Workie DL. Determinants of under-nutrition among children under five years of age in Ethiopia. *BMC Public Health*. 2020;20:399. <https://doi.org/10.1186/s12889-020-08539-2>.
111. Organization, W.H., Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group Joint Child Malnutrition Estimates: Key findings of the 2023 edition. 2023: World Health Organization.
112. Zewudie AT, Gelagay AA, Enyew EF. Determinants of Under-Five Child Mortality in Ethiopia: Analysis Using Ethiopian Demographic Health Survey, 2016. *Int J Pediatr*. 2020;2020(1):7471545.
113. Mayneris-Perxachs J, Swann JR. Metabolic phenotyping of malnutrition during the first 1000 days of life. *Eur J Nutr*. 2019;58(3):909–30.
114. Martins VJ, et al. Long-lasting effects of undernutrition. *Int J Environ Res Public Health*. 2011;8(6):1817–46.
115. Alderman H, Hoddinott J, Kinsey B. Long term consequences of early childhood malnutrition. *Oxf Econ Pap*. 2006;58(3):450–74.
116. Siami-Namini S, Tavakoli N, Namin AS. A comparison of ARIMA and LSTM in forecasting time series. In 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA). IEEE; 2018. p. 1394-1401.
117. Elsaraiti M, Merabet A. A comparative analysis of the arima and lstm predictive models and their effectiveness for predicting wind speed. *Energies*. 2021;14(20):6782.
118. Siami-Namini, S., N. Tavakoli, and A.S. Namin, A comparative analysis of forecasting financial time series using arima, lstm, and bilstm. *arXiv preprint arXiv:1911.09512*, 2019.
119. Pirani M, Thakkar P, Jivrani P, Bohara MH, Garg D. A comparative analysis of ARIMA, GRU, LSTM, and BiLSTM on financial time series forecasting. In 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE). IEEE; 2022. p. 1-6.
120. Goodfellow, I., *Deep learning*. 2016, MIT press.
121. Uzair M, Jamil N. Effects of hidden layers on the efficiency of neural networks. in 2020 IEEE 23rd international multitemic conference (INMIC). 2020. IEEE; p. 1–6.
122. Shen Z, Yang H, Zhang S. Neural network approximation: Three hidden layers are enough. *Neural Netw*. 2021;141:160–73.
123. Thomas AJ, et al. Two hidden layers are usually better than one. in *Engineering Applications of Neural Networks: 18th International Conference, EANN 2017, Athens, Greece, August 25–27, 2017, Proceedings*. Springer; 2017. p. 279–290.

124. Amin FM, Novitasari DCR. Identification of Stunting Disease using Anthropometry Data and Long Short-Term Memory (LSTM) Model. *Comp Eng Appl J.* 2022;11(1):25–36.
125. Kiser AC, Eilbeck K, Bucher BT. Developing an LSTM model to identify surgical site infections using electronic healthcare records. *AMIA Summits on Translational Science Proceedings.* 2023;2023:330.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.