




# Potato Consumption Forecasting Based on a Hybrid Stacked Deep Learning Model

Marwa Eed<sup>1</sup> · Amel Ali Alhussan<sup>2</sup> · Al-Seyday T. Qenawy<sup>3</sup> ·  
Ahmed M. Osman<sup>4</sup> · Ahmed M. Elshewey<sup>5</sup>  · Reham Arnous<sup>6</sup>



Received: 28 May 2024 / Accepted: 25 June 2024  
© The Author(s) 2024

## Abstract

Potato consumption forecasting is crucial for several stakeholders in the food market. Due to the market flexibility, the farmers can manipulate the volumes planted for a given type of produce to reduce costs and improve revenue. Consequently, it means that establishing optimal inventories or inventory levels is possible and critical in that sense for the sellers to avoid either inadequate inventory or excessive inventories that may lead to wastage. In addition, governments can predict future food deficits and put measures in place to guarantee that they have a steady supply of food some of the time, especially in regions that involve the use of potatoes. Increased potato-eating anticipation has advantages for the sellers and buyers of the potatoes. The experiments of this study employed various machine learning and deep learning (DL) models that comprise stacked long short-term memory (Stacked LSTM), convolutional neural network (CNN), random forest (RF), support vector regressor (SVR),  $K$ -nearest neighbour regressor (KNN), bagging regressor (BR), and dummy regressor (DR). During the study, it was discovered that the Stacked LSTM model had superior performance compared to the other models. The Stacked LSTM model achieved a mean squared error (MSE) of 0.0081, a mean absolute error (MAE) of 0.0801, a median absolute error (MedAE) of 0.0755, and a coefficient of determination ( $R^2$ ) value of 98.90%. These results demonstrate that our algorithms can reliably forecast global potato consumption until the year 2030.

**Keywords** Deep learning in agriculture · Forecasting · Long short-term memory · Potato consumption

## Introduction

The potato, recognized for its endless uses and high level of nutrients, is an important ingredient in recipes everywhere. It is the third most eaten crop, following wheat and rice. It has been eaten by over a billion people as part of their daily meals. The remarkable production statistics demonstrate the global need for potatoes,

---

Extended author information available on the last page of the article

which hit 375 million tonnes annually. The mean daily intake amounts to roughly 0.43 pounds per individual, equal to a medium-sized potato's measurement ('Potato Consumption per Capita,' n.d.). Potatoes are a widely used dietary source around the globe. Their offerings comprise a broad range of vital components, such as carbs, dietary fiber, vitamins, and minerals involving magnesium, potassium, and iron. Potatoes have a protein value equivalent to that of milk and fish, and it is higher than the amount of protein grade discovered in the wheat grains. In addition, potatoes provide an easily digestible energy source in comparison to grains such as rice and wheat. The potato is recognized for its low CO<sub>2</sub> imprint and low land usage compared to other crops. In addition, it includes the most important nutrients. Potatoes provide amino acids that are not found in other foods (Gustavsen 2021). The International Potato Center (2019) reports that more than 4000 potato varieties may grow well in many more environments than other plants. In 2008, the FAO admitted potatoes as a future food crop and a key source of nutritional stability, announcing it as the Year of the Potato.

After even more investigation, Eastern Europe is recognized for its significant preference for this carbohydrate-rich food. Belarus has the highest consumption rates, with its population consuming an impressive average of 155 kg per person yearly. In contrast, levels of consumption across some African countries are significantly lower. The yearly consumption rate in South Sudan is as low as 0.13 kg. These variations originated from a relationship between social norms and surrounding factors. Cultural factors play a significant role, as many regions have a rich history of using potatoes in their favourite dishes. Eastern Europe has considered potatoes an essential component of their diet, in addition to the fact that some African countries may have adopted potatoes into their cooking traditions at a later stage, leading to less advanced culinary practices associated with their utilization ('Potato Consumption by Country 2024,' n.d.).

Environmental conditions are critical. Potato growth takes place in moderate rainy regions, making climate adaptation an important aspect. Planting potatoes can introduce challenges in areas characterized by higher temperatures or a low water supply, which might directly affect overall consumption levels. The whole consumption of potatoes is affected worldwide because of some dietary habits caused by cultural and environmental restrictions. Eastern Europe has a superior performance in personal consumption, whereas other regions have a lower need for this food. Analysing these inconsistencies offers a more thorough understanding of the complex factors influencing our food.

Artificial intelligence (AI) supports the agricultural industry by bringing about change. Many technologies that fall under AI are used in agriculture, including supporting methods such as machine learning, computer vision, robotics, interactive systems, bio-inspired computing, and swarm intelligence. Farmers use AI tools to make decision-making regarding the usage of resources to improve productivity in farming. The machine learning (ML) algorithms use perceptions from possible sources such as sensors, satellites, and weather stations to grasp affairs concerning the quality of the soil and the health of crops, amongst other things concerning pest episodes and the like. This allows for the avoidance of wastage of water, fertilizers, and pesticides, making production more accessible. Using

'classical' approaches is less efficient than using machine learning models to anticipate crop yields. AI also assists the farmers in case there are some errors likely to happen in production, and it assists the farmers in a change by focusing on some of the practices through information on their performance and weather. Who needs traditional disease and pest sample scouts when you have facial recognition like picture identification for field diseases and pests from drone shots or a simple picture taken with a smartphone? Through it, farmers can prevent this from happening or use desired treatment measures to minimize the losses. Intelligent electronic devices are relatively wearables fitted with artificial intelligence needed to monitor the animals' health, their weights, and early signs of diseases to prevent. They ask robots to perform weeding, plant fruits and vegetables, and even operate tractors. They can be of great value in helping rectify the current shortage of skilled labor in agriculture and improving the productivity levels of the sector.

Machine learning is one of the smartest techniques that offers the analysis of data on consumption and identifying trends to forecast. The number of machine learning strategies for potato consumption predictions is high. Improved accuracy says that ML approaches can easily establish the relationship between potato consumption and other factors, which can be translated to more accurate forecasting than traditional methods. As indicated by data insights, ML models can identify the hidden structures, patterns, or trends that human analysts may not be able to visualize through data analysis for massive amounts of data. This can help in strategizing the way materials are produced, when to produce them, when to acquire raw materials, and when to order added materials for the current production process. Predictive data states that better predictions allow the stakeholders to forecast the possible future demand, achieve better resource allocation, and avoid having too much or insufficient product in stock.

A few of the traditional ML models merely make sense for potato consumption. A time series system also analyses past consumption figures to identify patterns and cyclical patterns over time. Some of these models are ARIMA, Auto-Regressive Integrated Moving Average, and SARIMA: Seasonal Auto-Regressive Integrated Moving Average. Linear regression models analyse the connections between potato consumption and other variables related to potato intake, including but not limited to income level, population demographics, and potato prices. The most common techniques include linear machines and support vector machines. The Long Short-Term Memory (LSTM) structure is one of the most efficient learning algorithms capable of addressing several facets of knowledge. In the case of sequential data, LSTMs are very good at approximating long-distance dependencies. Thus, many factors define the reliability of the machine learning forecasting models, including the quantity and quality of the available data initially used for training the model. The consumption levels of the past years act as the basis for the model. Information can be obtained from governmental bodies, agricultural departments, associations, or market research firms. The overall functionality of the model can be enhanced by incorporating an understanding of demography, which includes the population size and its evolution, income, food price, customer demand and weather, and the factors that have already been discussed.

Various researchers have demonstrated that using ML models helps predict the yield of potatoes. Gómez-Vegà et al. explored the potential of ML algorithms and traditional time series models to predict potato production in South Asia. The results revealed that the performance of ML models, particularly LSTM models, performed better compared to the traditional methods in terms of accuracy. Towfek and Alhusan (2024) documented the optimization of LSTM architecture for potato yield estimation. They made significant contributions using this model by incorporating the Balance Dynamic Biruni Earth Radius (BDBER) technique in the model to enhance model precision and robustness.

Several studies have demonstrated how ML models contribute to potato yield forecasts. Gómez-Vegà et al. (2019) explore the potential of ML algorithms and traditional time series models to predict potato production in South Asia. The results revealed that the performance of ML models, particularly LSTM models, performed better compared to the traditional methods in terms of accuracy. Towfek and Alhusan (2024) documented the optimization of LSTM architecture for potato yield estimation. They made significant contributions using this model by incorporating the Balance Dynamic Biruni Earth Radius (BDBER) technique in the model to enhance model precision and robustness.

## Problem Statement

However, it is worth noting that potato consumption may be predicted on a global scale. Potatoes are a major food crop, but their consumption system is not as simple as that of raw material crops such as rice or wheat. It is challenging to predict potato consumption in general. The intake of potatoes also varies across various countries ('Potato Consumption by Country 2024,' n.d.). The adoption of cultural symbols in traditional practices is also important. The needs of consumers will always be changing, with dietary trends that are focused on health and sustainability becoming a stronger consideration for consumers when it comes to their food choices. Low-carb diets and the growth of other potatoes' alternative crops, like cassava, may affect the consumption of potatoes. These variables are constantly changing, and models meant for forecasting must be flexible enough to account for these changes.

Factors such as temperature and precipitation affect the production of potatoes. Natural disasters or fluctuations in climatic conditions also impact the production process and are powerful tools for regulating the total supply and consumption. The data and potential disruptions that will be forecasted must suit the intended models of weather forecasting. Usually, obtaining accurate statistics on actual consumption levels of potatoes in many nations is not always possible. Lack of uniform data raising for different nations might hamper the formation of forecasting systems. Prediction for potato production will probably be one of the forecasts the stakeholders use in various aspects. To hold waste and guarantee the correct provision of products, farmers can utilize forecast data and change planting methods and methods for using resources. The information may also be helpful to food distributors and retailers in their inventory management and pricing policies; additionally, they can

ascertain whether there are enough potatoes to supply the consumer market without excess that leads to spoils. This information is helpful to governments in formulating policies on food safety and developing the agricultural sector, especially in areas where potatoes are famous. Potato industry players may be able to optimize presented aspects, reduce salvages, and guarantee food availability for a constantly increasing world population through participating or enhancing existing issues and prediction systems.

## Objectives

This paper demonstrates how ML can achieve several objectives for forecasting potato consumption. The ultimate objective is to enhance the effectiveness of more conventional approaches in forecasting. Such knowledge could mean that they grasp the dynamics of the relationships connecting the consumption of potatoes and several other factors and might offer more precise projections than simple statistical considerations. From the accumulated better decisions and accurate information, the entities associated with the life cycle of the potato make decisions. It is a case for manufacturers, distributors, retailers, government, and ministries involved in policymaking. Stakeholders can plan production and storage to anticipate shortages or surplus production to attain an optimum quantity for the target markets. It is precious for industries because such highly advanced systems themselves can analyse vast amounts of data intensively and deeply delve into the consumption of potatoes to achieve the most efficient production of potatoes; they can search for new outlets for the consumption of potatoes and form new types of potatoes tailored to meet the needs of various markets. The level of accuracy in the potato market can be used for stabilizing and enhancing its efficiency since an accurate forecast for the future rise or fall in the supply of potatoes or the demand for potatoes can be made. The participants in the market can adjust their prices to decrease the fluctuations in the market and ensure producers have an acceptable rate of return. Applying ML forecasting for potato production could help enhance sustainability due to the increased ability to predict future market demands for potatoes so producers, water, and fertilizer, amongst other resources, can be used efficiently with minimal wastage and losses in the food chain.

## Contribution

The experiments applied different techniques for machine learning, including stacked long short-term memory (Stacked LSTM), convolutional neural network (CNN), random forest, support vector machine,  $K$ -nearest neighbour, bagging regressor, and dummy regressor. The data collection (dataset) includes data about potato consumption over 60 years, starting from 1961 to ending in 2019. During the studies, it was discovered that the Stacked LSTM model had superior performance in comparison to the other models. The Stacked LSTM model achieved a mean squared error of 0.0081, a mean absolute error of 0.0801, a median absolute

error of 0.0755, and an  $R^2$  value of 98.90%. The study utilizes prediction evaluation techniques such as mean squared error, mean absolute error, median absolute error, and  $R^2$ . These results demonstrate that our algorithms could reliably forecast global potato consumption for the past 6 years and up until the year 2030.

The research article is structured as follows: The second section outlines previous studies regarding potato consumption forecasting. The third section introduces the methodology for data preparation, preprocessing methods, model setup, different ML, and performance evaluation metrics. The fourth section assesses the methodology with different models for potato consumption forecasting. At the end, the conclusion and future research direction are presented.

## Related Work

As the quantity of potatoes consumed in precise detail is of immense importance to stakeholders, particularly in the agricultural sector, it is imperative to get closer estimates of the future consumption of potatoes. In this respect, farmers can utilize it to develop better methods in crop farming and rearing stock for improved food productivity, as well as storage and marketing infrastructure for food that can prevent losses and increase farmers' earnings. On the same note, it enables the formulation of policies that track products traded in and out of the country to ensure the availability of the products that have been capacity-developed to support food security. This is very important for predicting the number of people consuming potatoes in the future as the analysis on trends through AI systems, mainly through machine learning methodologies, yields better results than conventional methods. Other part discusses studies where the use of AI-based techniques is presented or where there is an indication of potentially suitable applications for predicting potato consumption. It is focused on the types of models selected as well as on the ability to address the interdependencies that exist between latent factors and consumption patterns.

There have been several attempts to project the Indian production. Sahu et al. (2024) used various methods like data envelopment analysis, descriptive statistical analysis, price dependency analysis, and sustainability evaluation and several data sources to explore the potato industry of India from 1966 to 2020. Potato yields were found to be limited and erratic; prices were unstable, and inefficient markets failed to connect producers to the marketplace while export opportunities were dwindling in India. India was recommended to increase its potato production and marketing capabilities, export more efficiently using better methods, and increase its level of cooperation. Amongst the issues they discussed was contract farming with PepsiCo, where they improved the quality and quantity of production and secured loans, inputs, technical support, and guaranteed prices for their produce. They also pointed out other issues, such as failure to honour contractual terms, unfair treatment of farmers, and failure of the government to adequately regulate the industry. They proposed that the government should come in to regulate contract farming to ensure that contracts are fair and the farmers get protected, as well as, importantly, to make the sector more competitive and sustainable.

It was possible to determine optimal models for each country and obtain estimates of the output. Issues of weather conditions, insects, diseases, and other related factors explain the demand, consumption, trade, and production in the coffee market. The research was found to be of immense use in understanding the potential future supply of potatoes in South Asia that could be of the utmost significance in nutrition and farming.

Sahu (2006) forecasted the crop yield of potato, mustard, rice, and wheat data during the period 1971–2001 using AR and ARIMA, considering the effects of the irrigation method. Srivastava (2010) analysed potato yield in the Bareilly district of Uttar Pradesh with the help of fuzzy time series data and models with artificial neural networks. Kumar (2016) applied time series data from 1980 to 2015 to calculate the production and supply of potatoes in India. Analysis of the findings revealed that ARIMA (1,1,2) and ARIMA (2,1,3) best explained the probability of forecasting production and cultivation area. Singh et al. (2020) employed the biometric parameters of potatoes collected pre-harvested to develop prediction models for predicting yields of potatoes in Manipur.

The image segmentation method (Sentinel 2) developed by Ashourloo et al. (2020) is a new detection algorithm for potato fields. This method makes use of the fact that potato reflectance behaves differently through the various physical characteristics of light (spectral characteristics) and at different stages in the growth of potato (temporal characteristics). In the data acquisition process, the team obtained the following data: crop types from various sites were obtained at three locations in Iran (utilizing GPS from 2019 potato and other crop fields) and one location in the United States (from the Cropland Data Layer). To do this, they examined the reflectance pattern properties of different crops and focused on a close examination of the reflectance properties of potatoes; they were then able to identify four spectral features that were conspicuous in potatoes. They further found an approach to combine these features and then used it to train a system to recognize potato fields. The model attained over 90% in recognition of potato fields in all four studied districts. The findings of this research provide a convincing argument for the effectiveness of the use of high-resolution satellite data to map potato fields automatically and accurately.

Suh and colleagues (Suh et al. 2018) aimed to analyse whether a transfer learning approach that employed convolutional neural networks (CNNs) was effective for image classification. The researchers played three alternative versions of AlexNet in addition to the six conventional frameworks that included Inception-v3, VGG-19, GoogLeNet, ResNet-50, and ResNet-101. All networks were trained on the ImageNet dataset as initial training. Thereafter, they evaluated their performance by coming up with farm images of sugar beets and volunteer potatoes using different classifications in images under varying lighting conditions. The aggregator was able to generate the best prediction with 98% accuracy using a modified AlexNet architecture with two output units for a binary classification task. However, the performance of a modified VGG-19 also came up with strong accuracy (98.7%) in this binary classification problem. The results highlight the versatility of an approach based on transfer learning on various graphs, even for images captured in outdoor environments and having difficulty

adjusting to changes in illumination and background. Real-time applicability is usually item-specific (depending on the network used) and depends on the system design. Although the classification is fast and only takes less than 0.1s, a system implementation that would cover all courses could be required to sustain the pace throughout the whole training. The author then explains the importance of transfer learning for real-time weed detection in agricultural environments that require appropriately fast and accurate classification.

Yadav et al. (2024) presented several models for potato production predictions in the main Indian states. They compared ARIMA, a forecasting approach, and state space and XGBoost processes. These studies highlighted that there is not a right strategy approach at all. Rather, the most optimal model differs according to the data types of each state. The treatment of data characteristics and validation is equally essential to ensure predictions can be validly produced. One of the researchers' contributions is emphasizing the need to deal with the data in a real-world setting, including non-stationarity, significant outliers, and non-linearity. They suggest focusing on the fact that the points of understanding data complexities and each model's caveats are crucial to making reliable forecasts. Companies like Dole in India may find such quantitative models useful for streamlining agricultural policies and decisions based on the country's current potato consumption trends. Khorramifar et al. (2021) conducted an experiment to establish the ability of nine sensors to identify various VOCs present at different concentrations of 1–10,000 ppm. They used the data collected through sensors by applying machine learning techniques like PCA, LDA, and ANN. PCA resulted in the identification of the most significant and representative components that could account for almost 93% of the variance in the data. In addition, LDA and ANN exhibited high accuracy in VOC detection—100 and 96%, respectively.

The following article by Amkor and El Barbri (2023) describes operating a multimodal system comprising five MOX gas sensors algorithmically trained for the classification and prediction of the quality of potatoes fertilized with nitrogen, phosphorous, and potassium fertilizers or processed manure extracted from sheep and donkeys. The study classified information using the *K*-nearest-neighbour model with a 90% success rate. Developers used the NARX network for the forecasting task for which the accuracy was 0. The first algorithm had a correlation coefficient of 0.99 and produced minimal error. It was verified experimentally that the tool being presented was useful. Towfek and Alhussan (2024) show the application and effectiveness of the BDBER algorithm-trained LSTM models for short-term and long-term potato production prediction. They have scientifically established through empirical studies that the BDBER-tuned LSTM models outperform conventional algorithms in accuracy and prediction reliability. The flexibility of the BDBER algorithm in balancing the schedules of exploration and exploitation significantly enhances efficiency in using LSTM models for forecasting. The results affirm the importance of using artificial intelligence techniques in agricultural model development for better forecasting.

Boguszewska-Mańkowska et al. (2022) showed the use of ensemble learning models, and notably, the final classifier proved to be extreme gradient boosting. The last classifier was evaluated at 96.7% accuracy that incorporated several measured



potato parameters to discriminate amongst various tolerance groups. These parameters included relative yield decrease, stem mass, maturity, assimilation area, leaf mass, yield per plant, calculated climatic water balance, and indices the DSI and the MSTI. Elbasi et al. (2023) emphasize the use of machine learning algorithms and IoT sensors in contemporary agriculture to enhance crop production efficiency and minimize waste by making well-informed decisions. The study discusses obstacles and opportunities in combining these technologies and provides experimental results showing how altering labels affects data analysis algorithms. Studying data from farms, such as real-time IoT sensor data, can help farmers make better decisions on aspects that impact harvest growth. Although facing difficulties, the positive results suggest that ML methods will be essential for predicting future agricultural productivity. The study concluded that precise feature selection is crucial for enhancing the accuracy of ML systems. The maximum level of accuracy was attained by utilizing temperature, humidity, pH, and precipitation variables, resulting in 97.05% accuracy with Bayes Net and 97.32% accuracy with random forest. Advancements in this sector could enhance crop production efficiency, minimize waste, and enhance global food security.

Kumar and Patel (2023) introduced HDLCCNN, a method that improves the performance of the global potato crop. The strategy outperforms existing methods like VGG-INCEP, Deep CNN, random forest, and other spiking neural networks. The proposed method outperforms existing methods by 4% in accuracy, 6% in precision, 3% in recall, and 3.5% in *F*-score. It also offers higher specificity, sensitivity, and PSNR. Nazir et al. (2023) achieved a 98.12% accuracy in classifying different potato plant leaf illnesses (late blight, early blight, and healthy leaves) in 10,800 photos by utilizing the EfficientPNet tool. Experiments have proven that our method is effective for classifying potato plant leaf diseases and can handle distorted samples reliably. Therefore, farmers can reduce costs and increase yields by utilizing this technology.

In 2024, Sarah et al. conducted an experiment in which they trained three object detection models utilizing both conventional and super-resolved sets of data. Training on the super-resolved dataset improves the precision of identifying and locating a potato late blight disease. Faster-RCNN, Detr, and Yolo V8 outperformed the rest of the models in smoothly detecting small lesions on leaves with low training as well as inference time. The study by Mishra et al. (2024) assesses the efficacy of ARIMA and ETS models in forecasting potato output in India, China, and the USA. The study utilizes historical production data from these nations. The results indicate that the ETS model routinely surpasses the ARIMA model in forecasting potato production. China and India are projected to remain major contributors to the potato industry, further establishing themselves as essential participants.

Current research on predicting potato consumption with machine learning models is rare. Although researcher Yadav et al. (2024) discuss the economic aspects affecting potato demand, there is little emphasis on using machine learning to predict consumption. The integration of AI and ML into potato-related research has significant potential to enhance food security, increase the productivity of agriculture, and promote environmental sustainability to meet the rising need for this essential thing. Table 1 displays the comparison of the research mentioned above.

## Materials and Methods

The data used in the present study is about the consumption of potatoes for 60 years starting from 1961 to 2019, available at <https://www.kaggle.com/datasets/vishaljiotedra/potato-consumption-of-world-in-last-60-years>, and the data was divided into training and testing data sets where 80% of the data was used for training the model and 20% was used for testing the training model. These percentages imply that 80% of data was utilized in training the machine learning models and 20% was set aside for testing the models.

The reason behind using 80–20 proportion basis is clearly linked to the model testing and training aspect. An argument for using 80% for training is that the models would have ample data to learn from, which is important when attempting to identify pattern and trends in the data. This large training set is particularly helpful in developing models that are more general and accurate on new data that has not been used during the model creation.

On the other hand, using only 20% data for the test gives a realistic evaluation of the performance of the model. This testing set can be used to further validate the model on data that the model has not seen before and is a good way of gauging how well the model is going to perform in estimating the potato consumption in the future. The 80–20 split is a common practice in machine learning since it gives a reasonable number of examples for training the model and a reasonable number of samples for assessment of model's performance. Figure 1 shows the model representation for potato consumption forecasting.

The experiments applied different machine learning techniques, including Stacked LSTM, CNN, random forest, support vector machine,  $K$ -nearest neighbour, bagging regressor, and dummy regressor.

The model description for forecasting potato consumption is as follows:

1. The relevant data for potato consumption is determined as the input for the model. The dataset includes data about potato consumption over 60 years, from 1961 to 2019.
2. Normalization is an essential preprocessing step, mainly when dealing with approaches impacted by scaled features. We normalized our numerical features. This technique standardizes the scale of all features to prevent any one characteristic from overriding the others during model training.
3. The data were divided into 80% training and 20% testing.
4. The training data will be used for forecasting techniques in machine learning, including Stacked LSTM, CNN, random forest, support vector machine,  $K$ -nearest neighbour, bagging regressor, and dummy regressor.
5. Accuracy is evaluated using several metrics, such as mean squared error, mean absolute error, median absolute error, and  $R^2$ .
6. Potato consumption forecasting: depending on the evaluations, the more accurate forecasting model is selected to predict potato consumption in the future.

**Table 1** Summary of related works

Reference	Focus	Methodology and approach	Findings and outcomes
Sahu et al. (2024)	Potato industry in India	Descriptive statistics, sustainability evaluation, price dependence analysis	Low and inconsistent potato yields, volatile prices, poor market connections, diminishing exports
Sahu (2006)	Potato industry in India	Case study of contract farming with Pepsi	Improved production quality and profitability, potential exploitation of farmers
Kumar (2016)	Potato production trends in South Asia	Statistics and machine learning	Optimal models for each country, impact of weather and diseases on yield
Singh et al. (2020)	Potato, mustard, rice, wheat yield prediction in India	AR and ARIMA models	Irrigation impacts yield
Ashourloo et al. (2020)	Potato production and cultivation area estimation in India	ARIMA models	ARIMA (1,1,2) and (2,1,3) most accurate for production and area
Suh et al. (2018)	Potato yield estimation in Manipur	Biometrical characteristics	Developed pre-harvest prediction models
Yadav et al. (2024)	Potato field identification using satellite imagery	Spectral and temporal characteristics	Achieved high accuracy (over 90%) in identifying potato fields
Khorramifar et al. (2021)	Weed detection in sugar beets & potatoes	Convolutional neural networks (CNNs)	Achieved high accuracy (98%) in weed classification
Amkor and ElBarbri (2023)	Potato production forecasting in India	ARIMA, state space, XGBoost	No one-size-fits-all model, careful data analysis is crucial
Towfek and Alhussan (2024)	VOC detection for potato quality	Machine learning (PCA, LDA, ANN) on sensor data	Achieved high accuracy (100% and 96%) in VOC detection
Boguszewska-Manikowska et al. (2022)	Potato quality classification and prediction	Multimodal system with machine learning	Achieved 90% success rate in classification and 0.99 correlation coefficient in prediction
Elbassi et al. (2023)	Potato production prediction using LSTM models	BDBER-tuned LSTM models	Improved accuracy and reliability compared to conventional algorithms
	Potato tolerance group classification	Ensemble learning models	Achieved 96.7% accuracy using various measured potato parameters
	Crop production efficiency using machine learning and IoT	Machine learning algorithms on sensor data	Achieved high accuracy (97%) using temperature, humidity, pH, precipitation data

Table 1 (continued)

Reference	Focus	Methodology and approach	Findings and outcomes
Kumar and Patel (2023)	Global potato crop performance improvement	Deep learning (HDLCNN)	Outperformed existing methods in accuracy, precision, recall, <i>F</i> -score
Nazir et al. (2023)	Potato plant leaf disease classification	Efficient PNet tool	Achieved 98.12% accuracy in classifying potato diseases
Sarah et al. (2024)	Potato late blight detection using super-resolved datasets	Object detection models (Faster-RCNN, Detr, Yolo V8)	Improved accuracy in detecting potato late blight disease
Mishra et al. (2024)	Potato production forecasting in multiple countries	ARIMA and ETS models	ETS model outperforms ARIMA for forecasting potato production

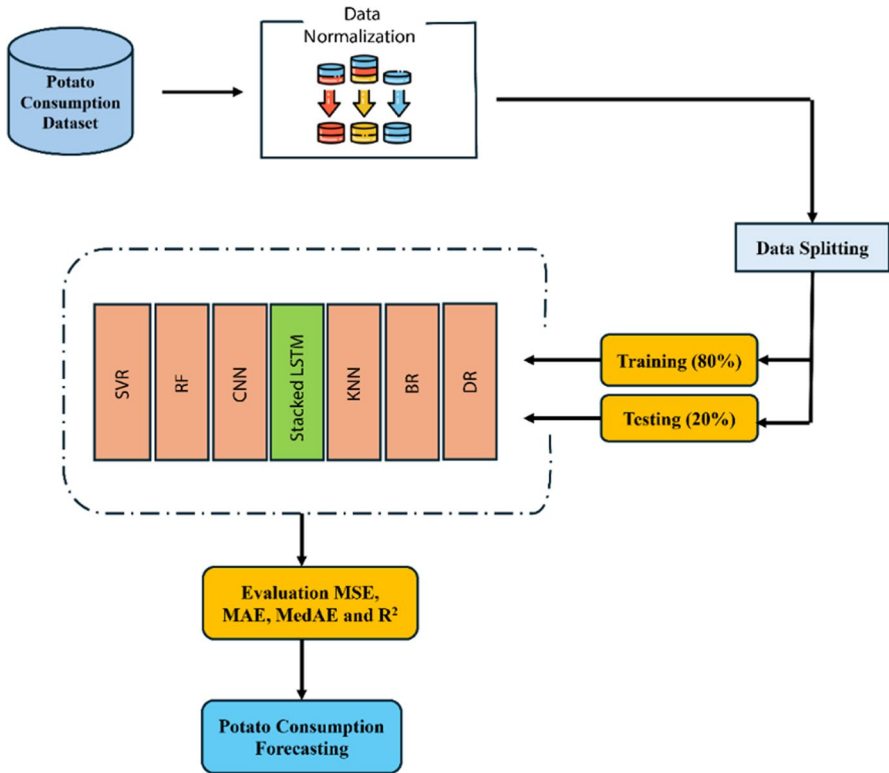


Fig. 1 The model represents global potato consumption until 2030

## Machine Learning (ML) and Deep Learning (DL) Models

### Long Short-Term Memory

LSTM is suitable for regression tasks as it is commonly applied to reasoning sequential data and reduces the significant correlations that arise from non-linear regression tasks. While classical statistical methodologies were built to make predictions over the immediate past of the assumed stochastic process, LSTMs can recall information about events even if these happen a long time after the actual query. This ability makes the model extremely effective at determining the likely occurrence of events in the distant future (Hochreiter & Schmidhuber 1997). This function is handy when performing regression predictions using time series data since the latter provides details of earlier data pointers for the time series. Experimental studies have shown applications of classical LSTM methods for regression problems in financial, traffic, or weather forecasting to be more effective when compared to classical methods (Graves et al. 2013)—the significance of LSTMs for regression of time-sequential data: best practice analyses and design changes.

### The Stacked Long Short-term Memory

The artificial neural network library used for the selected algorithm is LSTM = the stacked short-term memory network for this paper LSTM. Stacked LSTMs introduce the ability to build compound structures using the multiple multiplied copies of data passed through individual LSTM layers. Historical data on raw potato ingestion will be used as input to LSTM to obtain periodicity. This enables the model to predict the total and anticipated demand for fuel within a certain period, either in the short or long run. They also added stacked LSTMs with several parameters to obtain a better prediction of potato consumption. The model contains three layers: how many layers of hidden units does the network have? 128; 64; 32; 14 batch size = 32 15; learning rate = 0. The learning rate is 0.0001 and momentum 0.9 with epochs at 50, the optimizer is the Adam type, the time step is 32, and the output activation is linear. The structural view of the model is shown in Fig. 2.

### Convolutional Neural Network (CNN)

While the long short-term memory (LSTM) network emerged as the most successful model, discussing the rationale behind including a convolutional neural network (CNN) in the experiment is valuable. While seemingly counterintuitive for time series data like potato consumption, CNNs possess capabilities. Although potato consumption is a global phenomenon, there might be underlying spatial biases or regional variations in consumption patterns. For instance, climate or cultural

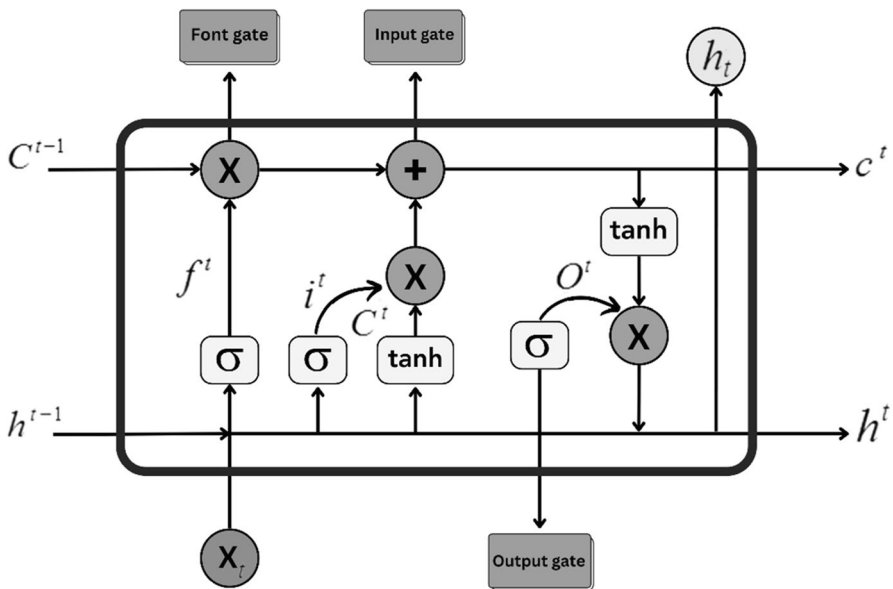


Fig. 2 Long Short-term Memory structure

preferences could influence consumption levels in specific geographical areas (El-Kenawy et al. 2023).

There are potential uses of CNNs, such as if potato consumption data was structured as a grid (e.g. average consumption per country), a CNN could extract relevant spatial features from this grid-like representation. These features might capture regional consumption patterns or relationships between neighbouring countries. CNNs can be effective at identifying patterns within images. If consumption data exhibited any visual trends over time that could be represented spatially (e.g. heat-maps), a CNN might be able to learn these trends and use them for forecasting.

## **Random Forest**

The random forest regressor model was the second noisiest model, but it provided more insights than the other models, except the LSTM model. This section further explains the central ideas associated with random forests and how applicable they can be in regression issues, such as estimating the amount of potatoes and individual uses. Random forest regression is a supervised learning approach that uses individual decision trees for predictions that consist of several decision trees. It is executed on randomly selected features (factors affecting potato consumption) and a randomly collected part of the records from the original dataset. It also reduces bias and helps decrease the model's overfitting of the training data. Random forests have several strengths that can be utilized for forecasting potato consumption. Several factors can impact the consumption of potatoes, such as population growth, the current economic stand, and popularity. Random forests enable complex relations within the data without requiring feature engineering. Implicit coding: Real-world data often includes noise and irrelevant observations. Classification by random forests is practically not prone to anomalies since the latter can be effectively averaged out using multiple decision trees. This is important when reconciling what took place in the past, given that there may be discrepancies regarding the consumption of potatoes. Such models are comparatively easy to read compared to some complex models, such as random forests. It is possible to determine the features most impacting these levels by examining the values of an individual feature in the importance score.

## **Support Vector Machine (SVM)**

Support vector machines (SVMs) discover a hyperplane in a vector space of a higher dimension than the actual dimension in which the data points are defined to be well separable and well apart from the hyperplane. This margin is the decision boundary for regression, where we try to minimize the cost derived from the prediction error for future data. Neither a positive correlation between population changes or variables of economic indicators and potato consumption is demonstrated. SVMs can also account for the non-linearity of these relationships using a kernel function to map the data into a higher dimensional feature space that captures this non-linear correlation. There is a possibility that the regression models might be affected by the occurrence of outliers in the dataset. SVMs are inherently less prone to outliers

because they give most of their attention to the data points along the plane representing a higher number of classifications termed support vectors. These metrics signify that when additional features, such as consumption data, are added to the models, the outcome depends on several factors. SVMs can deal with feature spaces of higher dimensionality through kernel functions because they can easily handle higher-dimensional data with ease by using the kernel function.

While the experiment highlighted LSTM's superiority, SVM regression remains a viable option for potato consumption forecasting. Its ability to handle non-linear relationships, deal with outliers, and work with high-dimensional data makes it a valuable tool. However, kernel selection and interpretability require careful consideration. Future research could compare different kernel functions and explore techniques to improve the interpretability of SVM predictions.

### **K-Nearest Neighbour (KNN)**

KNN is a non-parametric machine learning approach for regression tasks. It makes predictions based on the similarity of new data points to existing data points in the training set. Its mechanism considers many aspects. Like Euclidean distance, a distance metric calculates the similarity between the new data point (an unknown consumption year) and all data points in the training set (past years with known consumption). A predefined number of nearest neighbours ( $K$ ) are identified based on the calculated distances. In regression,  $K$  is typically a small, odd number. The average consumption value of the  $K$ -nearest neighbours becomes the predicted consumption value for the new data point (Zaki et al. 2023).

While KNN offers simplicity and interpretability, its limitations likely contributed to its lower performance than LSTM in this specific potato consumption regression task. However, KNN can still be a valuable tool for initial exploration and gaining essential insights into regression problems, particularly with well-behaved, low-dimensional datasets.

### **Bagging Regressor**

Bagging is a technique or type of ensemble learning used to enhance the stability and accuracy of the decision tree prediction model. It does this by creating several copies of a particular base model, each trained on different samples of the original data. This way, each base learner might concentrate on different aspects of the data, which reduces the variance and leads to a more stable final model.

When predicting future potato consumption, some natural random elements might be involved in the data. Different factors might affect consumption at a given time in a year, and a particular model might be susceptible to these changes. In this context, a bagging classifier trains multiple base learners on various data groups from the dataset. Then, it makes a collective prediction by eliminating the reliance on individual data points. This makes it easier to provide a more accurate prediction and thereby helps minimize the effects of noise in the data. Such a parallel approach ensures that each base learner is trained on



a slightly different portion of the data; naturally, the final model obtained will have more natural coverage over the variability in potato consumption. This has the effect of enabling better forecasting of future consumption of new samples of the collected data. The bagging techniques are instrumental in estimating the uncertainty and predicting the outcome. Prediction results from each learner can be used to construct a visualization of the confidence of the full model outputs through the spread of predictions.

Bagging represents an effective learning scheme that implements ensemble learning in regression tasks. Weather forecasting for potato consumption can contribute to a better understanding of the issues related to data variability and assist in obtaining reliable predictions. Although the LSTM provided better results for this experiment, the bagging technique remains a useful predictor tool for regression analysis, including situations with potentially noisy or stochastic variables.

### Dummy Regressor

A dummy regressor is another basic machine learning model in which the foreseen outputs are made according to a specific plan, ignoring the actual data of the inputs. The implementation of a dummy regressor can be split into two primary cases. The first is the constant prediction, which means the model predicts that all the points will equal some value. For example, the dummy regressor regarding potatoes might continually signify the average consumption level for the past 60 years. The second one is the majority prediction, which is more common in classification problems, wherein a dummy regressor predicts the most used class in the training set.

The dummy regressor is used as a benchmark to examine the model's performance. In case the considered intricate model like the LSTM network is not much more beneficial than, for instance, the dummy regressor, it points to the fact that it might involve numerous unnecessary complexities that are not capable of adequately reflecting the relations in the data. Relative to a model of more complexity, one can check the contribution of each additional component by comparing it to a dummy regressor. This reveals that if the complex model does not generate significantly better results than the dummy regressors, the same result could have been obtained using a simpler model but with less computational efficiency. Comparison can also determine whether it is rational to have a complex model or if a simple one is good enough.

### Results and Discussion

Figure 3 depicts a 60-year trend (1961–2019) in global potato consumption. Consumption is measured in thousands of tonnes (*y*-axis) and plotted against years (*x*-axis). The graph suggests a generally upward trend, with some periods experiencing steeper increases. This indicates a relatively consistent rise in potato

consumption over time, with occasional bursts of more rapid growth. The overall picture suggests a positive development in global potato agriculture, possibly reflecting increased consumption, expanded cultivation areas, or growing demand driving production.

## Evaluation Metrics

The experiment leveraged Jupyter Notebook version 6.4.6, an interactive development environment, to run the machine learning models. This user-friendly platform simplifies Python code creation and execution. Jupyter Notebook operates within a web browser and allows compatibility with other programming languages, including Python 3.8. The hardware running the models consisted of an Intel Core i7 processor, 16 GB of RAM, and the Microsoft Windows 10 operating system.

The experiments applied different techniques for machine learning, including Stacked LSTM, CNN, random forest, support vector machine,  $K$ -nearest neighbour, bagging regressor, and dummy regressor. The collection of data (dataset) includes data about potato consumption over 60 years, starting in 1961 and ending in 2019. During the studies, it was discovered that the Stacked LSTM model had superior performance compared to the other models. The study utilizes prediction evaluation techniques such as mean squared error, mean absolute error, median absolute error, and  $R^2$  calculated through the following:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (\text{Act}_i - \text{pre}_i)^2$$

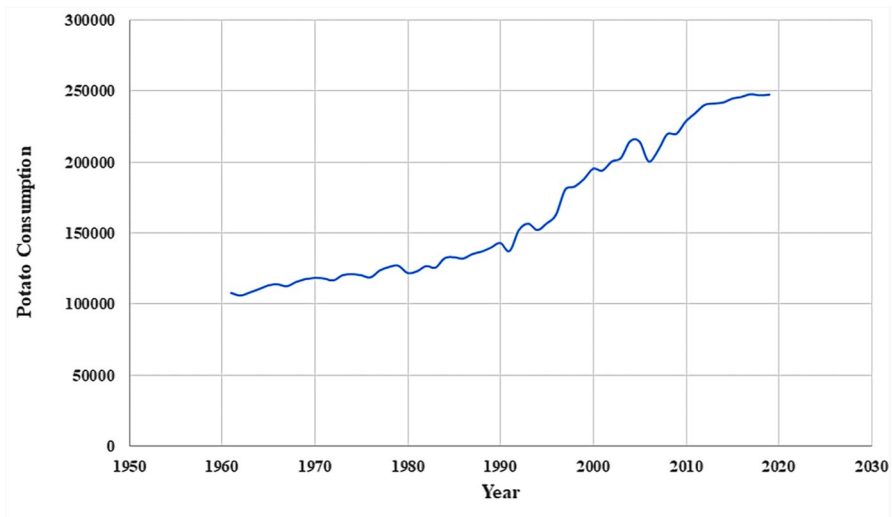


Fig. 3 Trend of potato consumption in global (1961–2019) in thousands of tonnes

$$MAE = \frac{1}{n} \sum_{i=1}^n |Act_i - pre_i|$$

$$MedAE = \text{median}(|pre_i - Act_1|, \dots, |pre_i - Act_i|)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Act_i - pre_i)^2}{\sum_{i=1}^n ((\sum_{i=1}^n Act_i) - Act_i)^2}$$

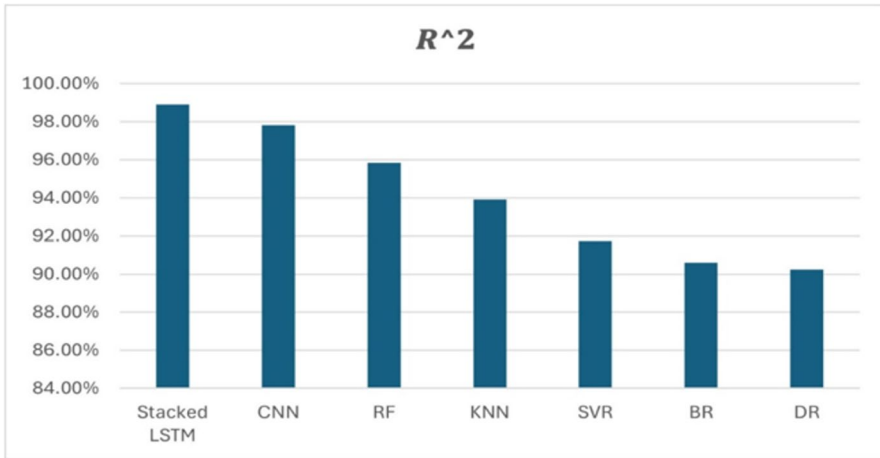
Table 2 displays the performance of different techniques for machine learning, including Stacked LSTM, CNN, random forest, support vector machine, *K*-nearest neighbour, bagging regressor, and dummy regressor. As shown in Table 2, the Stacked LSTM model achieved a mean squared error of 0.0081, a mean absolute error of 0.0801, a median absolute error of 0.0755, and an *R*<sup>2</sup> value of 98.90%, while the DR model achieved the lowest values: a mean squared error of 0.0326, a mean absolute error of 0.3594, a median absolute error of 0.3629, and an *R*<sup>2</sup> value of 90.25%. The CCN model’s MSE, MAE, MedAE, and *R*<sup>2</sup> are 0.0098, 0.0962, 0.0845, and 97.83%. The RF obtained an MSE of 0.0174, an MAE of 0.1531, an MedAE of 0.1482, and an *R*<sup>2</sup> of 95.83%. The KNN obtained an MSE of 0.0194, an MAE of 0.1843, an MedAE of 0.1958, and an *R*<sup>2</sup> of 93.91.83%. The SVR obtained an MSE of 0.0269, an MAE of 0.2841, an MedAE of 0.2735, and an *R*<sup>2</sup> of 91.74%. The BR obtained an MSE of 0.0292, an MAE of 0.3283, an MedAE of 0.3152, and an *R*<sup>2</sup> of 90.61%.

Figure 4 illustrates the coefficient of determination (*R*<sup>2</sup>) values for various regression models used to predict potato consumption. The models evaluated include the stacked long short-term memory (Stacked LSTM), convolutional neural network (CNN), random forest (RF), *K*-nearest neighbour (KNN), support vector regressor (SVR), bagging regressor (BR), and dummy regressor (DR).

The Stacked LSTM model achieved the highest *R*<sup>2</sup> value, close to 99%, indicating its superior performance in capturing the variance in the data and making accurate predictions. The CNN model follows with an *R*<sup>2</sup> value of approximately 98%, showcasing its strong performance as well. The random forest model achieved an *R*<sup>2</sup> value of around 96%, demonstrating its effectiveness but still trailing behind the Stacked LSTM and CNN models.

**Table 2** Performance of the proposed Stacked LSTM model and other individual regression models

Models	MSE	MAE	MedAE	<i>R</i> <sup>2</sup>
Stacked LSTM	0.008	0.08	0.0755	98.90%
CNN	0.01	0.096	0.0845	97.83%
Random forest	0.017	0.153	0.1482	95.83%
KNN	0.019	0.184	0.1958	93.91%
SVR	0.026	0.284	0.2735	91.74%
Bayesian ridge	0.029	0.328	0.3152	90.61%
Decision tree	0.033	0.359	0.3629	90.25%



**Fig. 4**  $R^2$  of the Stacked LSTM model and several regression models

The  $K$ -nearest neighbour model showed a notable  $R^2$  value of about 94%, while the support vector regressor model had an  $R^2$  value of around 92%, indicating moderate performance in comparison to the top models. The bagging regressor and dummy regressor models had the lowest  $R^2$  values, with the dummy regressor scoring just above 90%, serving as a baseline for comparison. The bagging regressor also performed similarly to the dummy regressor, indicating that these models were less effective in capturing the complex relationships in the data compared to the other more sophisticated models.

Figure 5 shows how the Stacked LSTM approach in predicting the potato consumption performed. The straight line on the scatter plot connects the actual and the predicted values where the  $x$ -axis represents actual values of potato consumption, and the  $y$ -axis represents the values predicted. In the scatter plot displayed below, the data points lie almost on the diagonal line, thus proving that there is a strong positive relationship between the actual and predicted values.

Such alignment reinforces the validity of the predictions made by the model and those source points that are closer to the diagonal are the best. The above results reveal that the Stacked LSTM model's capability to learn and decompose the given data's latent structure fully supports its phenomenon-better forecast of potato consumption.

In general, the visualization given by the scatter plot corresponds to the high accuracy of the model that was mentioned earlier in terms of quantitative measures. The proposed Stacked LSTM model holds potential as a viable tool for the prediction of future consumption of potato by consumers which in turn will provide useful information for players in the value chain particularly the growers in agriculture.

Figure 6 illustrates the comparison of mean squared error and mean absolute error against the number of epochs used to train and test the Stacked LSTM model for predicting potato consumption.

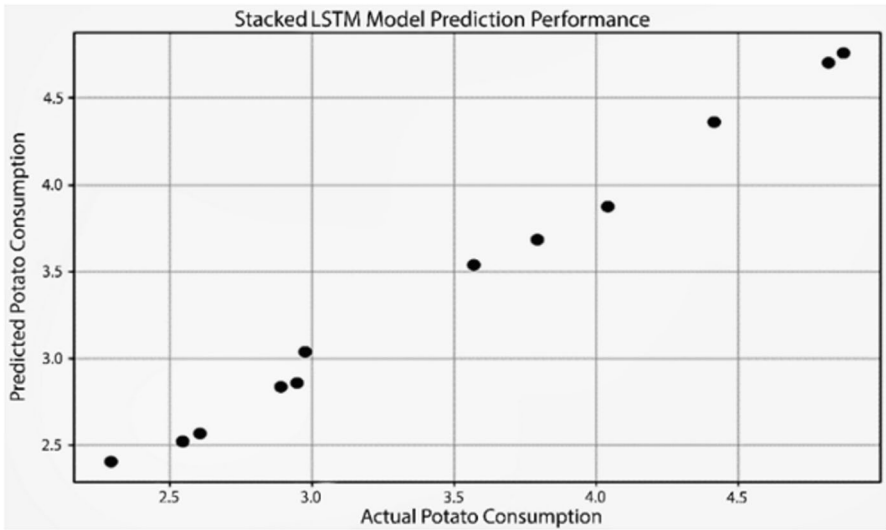


Fig. 5 Performance of the Stacked LSTM model

Figure 7 presents a visual representation of historical data and future forecasts for global potato consumption using the Stacked LSTM model. The x-axis represents the years, ranging from 1950 to 2040, while the y-axis represents potato consumption in thousands of tonnes. The blue dots illustrate the actual historical data on potato consumption from 1960 to 2019. The data shows a generally increasing trend in potato consumption over the years, with noticeable periods of rapid growth, particularly from the 1990s onwards.

The orange dots represent the model’s forecast for future potato consumption, extending from 2020 to 2030. The predictions indicate a continued upward trend in potato consumption, suggesting that demand for potatoes will keep rising in the coming years. This graph highlights the model’s ability to extend its learning from historical data to make reliable future predictions. The clear upward trajectory in

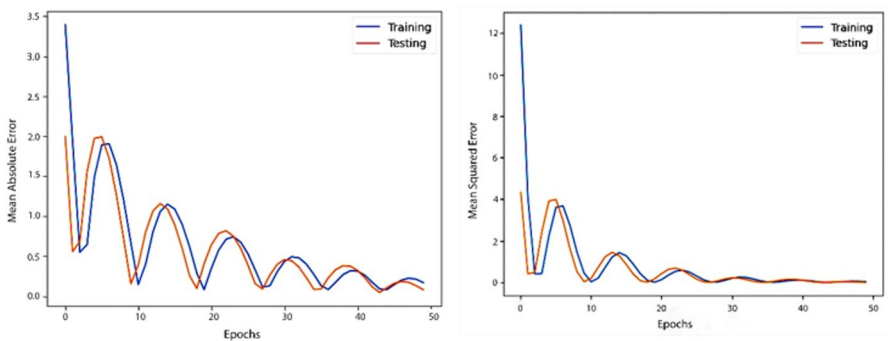
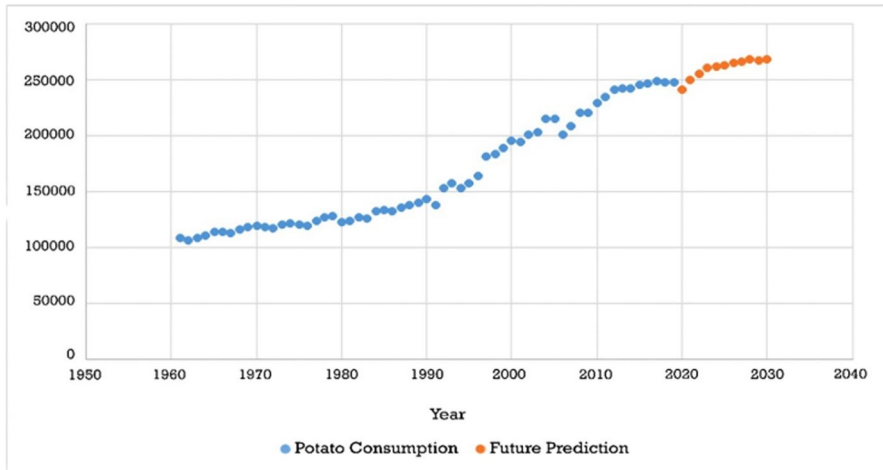


Fig. 6 Squared error and mean absolute error vs number of epochs using the Stacked LSTM model



**Fig. 7** Potato consumption forecasting (in thousands of tonnes) using the stacked LSTM model

both historical and forecasted data underscores the growing importance of potatoes in the global food market and the effectiveness of the Stacked LSTM model in capturing these trends for future planning and decision-making.

## Conclusion and Future Work

The experiments used machine learning methods such as Stacked LSTM, CNN, random tree, support vector, KNN, bagging regressor, and dummy regressor. The data set comprises records on the consumption of potato over a period of 60 years, from the year 1961 to the year 2019. In the course of the analysis on the models, they found out that the Stacked LSTM model had the best performance amongst the models. The Stacked LSTM model achieved a mean squared error of 0.0081, a mean absolute error of 0.0801, a median absolute error of 0.0755, and an  $R^2$  value of 98.90%. The study utilizes prediction evaluation techniques such as mean squared error, mean absolute error, median absolute error, and  $R^2$ . These data show that our algorithms can predict worldwide potato consumption till 2030.

It will be helpful for future research to cast AI techniques in the context of the forecasting models discussed above to comprehend the underlying forces behind the model. This could be very helpful in understanding the elemental interactions of potato consumption and add more value to the interpretation of the forecasts. However, this study aimed at the global level of potato consumption; further research could pinpoint various regional or subcategories of potato varieties. Thus, creating models that predict the usage of concrete potatoes for particular regions may help define the principles of efficient agricultural activity and the usage of available resources.

**Acknowledgements** Princess Nourah bint Abdulrahman University Researchers Supporting Project number PNURSP2024R 308, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Author Contribution** All authors have contributed equally.

**Funding** Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

**Data Availability** Data is available at <https://www.kaggle.com/datasets/vishaljiodedra/potato-consumption-of-world-in-last-60-years>.

## Declarations

**Ethics Approval and Consent to Participate** Not applicable.

**Consent for Publication** Not applicable.

**Conflict of Interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

## References

- Amkor A, El Barbri N (2023) Artificial intelligence methods for classification and prediction of potatoes harvested from fertilized soil based on a sensor array response. *Sens Actuators, A* 349:114106. <https://doi.org/10.1016/j.sna.2022.114106>
- Ashourloo D, Shahrabi HS, Azadbakht M, Rad AM, Aghighi H, Radiom S (2020) A novel method for automatic potato mapping using time series of Sentinel-2 images. *Comput Electron Agricult* 175:105583. <https://doi.org/10.1016/j.compag.2020.105583>
- Boguszewska-Mańkowska D, Ruszczak B, Zarzyńska K (2022) Classification of potato varieties drought stress tolerance using supervised learning. *Appl Sci* 12(4):1939. <https://doi.org/10.3390/app12041939>
- Elbasi E, Zaki C, Topcu AE, Abdelbaki W, Zreikat AI, Cina E, ... Saker L (2023) Crop prediction model using machine learning algorithms. *Appl Sci* 13(16):9288. <https://doi.org/10.3390/app13169288>
- El-Kenawy E-SM, Abdelhamid AA, Ibrahim A, Eid MM, Rizk FH, Zaki AM (2023) Time series forecasting of cryptocurrency prices with long short-term memory networks. *Fin Technol Innov* 2(2):18–26. <https://doi.org/10.54216/JSDGT.030205>
- FAO Food and Agriculture Organization of the United Nations (2008) International Year of the Potato. <https://www.fao.org/agriculture/crops/core-themes/theme/hort-indust-crops/international-year-of-the-potato/en/>. Accessed 12 May 2024
- Gómez D, Salvador P, Sanz J, Casanova JL (2019) Potato yield prediction using machine learning techniques and Sentinel 2 data. *Remote Sens* 11(15):1745. <https://doi.org/10.3390/rs11151745>
- Graves A, Mohamed AR, Hinton G (2013) Speech recognition with deep recurrent neural networks. 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. <https://doi.org/10.1109/icassp.2013.6638947>

- Gustavsen GW (2021) Sustainability and potato consumption. *Potato Res* 64(4):571–586. <https://doi.org/10.1007/s11540-021-09493-1>
- Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Comput* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- International Potato Center (2019) Potato: facts and figures. <https://cipotato.org/crops/potato/potato-facts-and-figures/>. Accessed May 12, 2024
- Khorrarifar A, Rasekh M, Karami H, Malaga-Toboła U, Gancarz M (2021) A machine learning method for classification and identification of potato cultivars based on the reaction of MOS type sensor-array. *Sensors* 21(17):5836. <https://doi.org/10.3390/s21175836>
- Kumar G (2016) Forecasting production and area of potato cultivation in India using ARIMA model. *Int J Soc Sci Develop Policy* 2(2):15–28
- Kumar A, Patel VK (2023) Classification and identification of disease in potato leaf using hierarchical based deep learning convolutional neural network. *Multimedia Tools Appl* 82(20):31101–31127. <https://doi.org/10.1007/s11042-023-14663-z>
- Mishra P, Alhussan AA, Khafaga DS, Lal P, Ray S, Abotaleb M, ... El-kenawy ESM (2024) Forecasting production of potato for a sustainable future: global market analysis. *Potato Res*. <https://doi.org/10.1007/s11540-024-09717-0>
- Nazir T, Iqbal MM, Jabbar S, Hussain A, Albathan M (2023) EfficientPNet—an optimized and efficient deep learning approach for classifying disease of potato plant leaves. *Agriculture* 13(4):841. <https://doi.org/10.3390/agriculture13040841>
- Potato Consumption by Country 2024. (n.d.) Retrieved from <https://www.kaggle.com/datasets/vishaljiode/potato-consumption-of-world-in-last-60-years>
- Potato consumption per capita. (n.d.). Retrieved from <https://www.helgilibrary.com/indicators/potato-consumption-per-capita>
- Sahu PK (2006) Forecasting yield behavior of potato, mustard, rice, and wheat under irrigation. *J Veg Sci* 12(1):81–99
- Sahu PK, Das M, Sarkar B et al (2024) Potato production in India: a critical appraisal on sustainability, forecasting, price and export behaviour. *Potato Res*. <https://doi.org/10.1007/s11540-023-09682-0>
- Sarah M, Abdlemadjid M, Sarah B, Yacine H, Miloud CEM (2024) Evaluating the effect of super-resolution for automatic plant disease detection: application to potato late blight detection. *Multimedia Tools Appl*. <https://doi.org/10.1007/s11042-024-18574-5>
- Singh NO, Singh NG, Singh LN, Kumar S, Paul AK (2020) Forecasting model of potato yield from farmers' fields in Manipur. *Int J Agricult Stat Sci* 16(1):401–405
- Srivastava RK (2010) Soft computing techniques: an application to short term forecast of potato production. *Int J Inform Technol Knowledge Manage* 2(2):205–210
- Suh HK, IJsselmuiden IJ, Hofstee JW, van Henten EJ (2018) Transfer learning for the classification of sugar beet and volunteer potato under field conditions. *Biosyst Eng* 174:50–65. <https://doi.org/10.1016/j.biosystemseng.2018.06.017>
- Towfek SK, Alhussan AA (2024) Potato production forecasting based on balance dynamic Biruni earth radius algorithm for long short-term memory models. *Potato Res*. <https://doi.org/10.1007/s11540-024-09721-4>
- Yadav S, Al khatib AMG, Alshaib BM, Ranjan S, Kumari B, Alkader NA, ... Kapoor P (2024) Decoding potato power: a global forecast of production with machine learning and state-of-the-art techniques. *Potato Res*. <https://doi.org/10.1007/s11540-024-09705-4>
- Zaki AM, Abdelhamid AA, Ibrahim A, Eid MM, El-Kenawy E-SM (2023) Enhancing K-nearest neighbors algorithm in wireless sensor networks through stochastic fractal search and particle swarm optimization. *J Cybersecurity Inform Manag* 13(1):76–84. <https://doi.org/10.54216/JCIM.130108>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



## Authors and Affiliations

**Marwa Eed<sup>1</sup> · Amel Ali Alhussan<sup>2</sup> · Al-Seyday T. Qenawy<sup>3</sup> · Ahmed M. Osman<sup>4</sup> · Ahmed M. Elshewey<sup>5</sup>  · Reham Arnous<sup>6</sup>**

✉ Ahmed M. Elshewey  
ahmed.elshewey@fci.suezuni.edu.eg

- <sup>1</sup> Faculty of Artificial Intelligence, Delta University for Science and Technology, Mansoura 11152, Egypt
- <sup>2</sup> Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, 11671 Riyadh, Saudi Arabia
- <sup>3</sup> Intelligent Systems and Machine Learning Lab, Shenzhen 518000, China
- <sup>4</sup> Department of Information Systems, Faculty of Computers and Information, Suez University, P.O. Box 43221, Suez, Egypt
- <sup>5</sup> Department of Computer Science, Faculty of Computers and Information, Suez University, P.O. Box 43221, Suez, Egypt
- <sup>6</sup> Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura 35111, Egypt