Sistemasi: Jurnal Sistem Informasi
Volume 14, Nomor 5, 2025: 2531-2539
ISSN:2302-8149
e-ISSN:2540-9719

A Hybrid Internet of Behavior Algorithm for Predicting IoT Data of Plant Growing using LSTM and NB Models

[1]Khansaa Yaseen*, [2]Omar Muayad Abdullah
[1,2]Department of Computer Science, College of Computer Sciences and Mathematics, University of Mosul, Iraq
*email: [1]khansaa.23csp64@student.uomosul.edu.iq,[2]omaraldewachy@uomosul.edu.iq

(**received**: 26 May 2025, **revised**: 21 July 2025, **accepted**: 22 July 2025)

Abstract

The researches that compare the accuracy between classical statistical prediction procedures and deep learning algorithms represent an important and modern field. The prediction accuracy of the plant growing is considered as an important factor in the field of smart agricultural technologies. This research proposes a hybrid Internet of Behaviors (IoB) technique that linking between time-series predicting and the classification models to estimate the plant growing behaviors using real environmental data. The proposed algorithm includes ML algorithms, especially Recurrents Neural Networks (RNN) and Long Short-Term Memory (LSTM), used for predicting the plant growing depending on sensor data. To improve the prediction accuracy, the outputs of the LSTM system were used as inputs to the Naïve Bayes algorithm. The dataset is collected from the Kaggle website using Internet of Things (IoT) sensor readings depending on the factors that affecting the plant growing. The obtained results stated that the proposed hybrid algorithm enhanced the prediction accuracy compared to using LSTM alone. Additionally, the using of Naïve Bayes algorithm added more reliable to the process of examining the growing behavior, making the proposed system more practical and provide the rapidity in task performing.

Keywords: *internet of behaviors (IoB), long-short term memory(LSTM), naïve bayes*

1 Introduction

The highspeed progress in Information Technology leads to a notable transformation in the way information is grouped, processed, and used among many areas [1]. Newly, the link between AI and the Internet of Things (IoT) has developed an idea of how digital techniques deal with human behaviors. Its activity has proved by activating the remotely communication devices to communicate and participate data in real times using the Internet. Recently, it is considered as an important technique for activating systems to make intelligent conclusions, which helps enhance the efficiency, minimize costs, and improve the user events [2]. The most important concept that integrate real time behavioral data monitoring is the Internet of Behaviors (IoB**)**, which is a platform that combines behavioral data grouped from IoT tools (such as sensors, smart communicating devices, and media platforms) with AI algorithms in order to analyze and estimate the behavior (Figure 1). IoB function is to analyze the user behavioral procedures by determining preferences, then improving their events. It also used in predicting the future behavior, so, IoB can be named as an advanced of the Internet of Things [3].

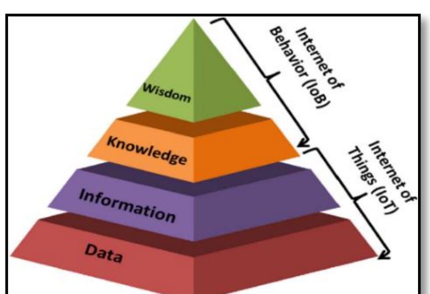

Figure 1 IoT and IoB diagram

 ISSN:2302-8149
e-ISSN:2540-9719

Agricultural sector has witnessed a remarkable transformation following the use of various artificial intelligence methods, which have helped improve crop management processes and rationalize resource use. [4]. For all significant progress, current research often does not address prediction and classification tasks together. Therefore, we have only found very few studies have combined these two approaches within the IoB framework to address agricultural prediction problems, which represents a clear research gap, especially in integrating behavioral data with predictive algorithms to achieve more accurate plant growth modeling. To determine this gap, the resented study tries to integrate the predictive capabilities of LSTM and the classification power of Naïve Bayes, based on real-time data collection, all within the framework of the Internet of Behaviors (IoB) to improve the accuracy of plant growth prediction while considering user behavior and agricultural decision-making.

LSTM networks are ideal for analyzing time-dependent data. An LSTM layer contains multiple recurrent memory objects (each consisting of else one or many of memory slots, each with three components: inputs, outputs, and forget, which are used to run operations such as reading, writing, and resetting) [5][6]. So, LSTM can determine long-term factors used in sequential data [7], and this procedure is considered suitable for predicting plant growing over time depending on the physical and environmental used variables.

In comparison, the Naïve Bayes technique is a clear yet efficient classifier that is capable of dealing with large datasets. According to its simple structure, many of statistical tests can be executed using this model [8]. simultaneously, when multiple features analyzed by Naïve Bayes, it considers each feature as a contributor, this means that the existence (or not) of one characteristic of a class does not affect the existence (or not) for other characteristics [9]. It is named “naive” because it supposes that all characteristics are independent of each other, which means may not always be right in real data of the world [10]. By applying Naïve Bayes after LSTM, the model can improve the prediction accuracy and classify the growing behavior more efficiently, and help in improving the decision-made in smart agriculture.

2 Literature Review

Several papers have used the term "Internet of Behaviors (IoB)" from different perspectives (Table 1). A paper introduced by Javaid et al, (2021) named *" Internet of Behaviors (IoB) and its role in customer services"*, the authors estimated how IoB techniques can help in analyzing customer behaviors and improve their convictions through tracing their choices. This is done by gathering data from many references like websites, cellular devices, and then analyzing it. The paper stated that IoB techniques are capable of determine the customer convictions and rearrange them though gathering the data based on their uses and behaviors. This enables the prediction of the future decisions [3]. Also, the research presented by (Stary, C., 2020) named *"The Internet-of-Behaviors as organizational transformations space with choreographic Intelligence"*, stated that designing IoT systems based on behavioral concepts is a more complicated field. Therefore, it needs more adjustable technical capabilities that match to users’ behaviors and, determining the future drawbacks in constructing intelligent systems based on the behavioral concepts [11].

The researchers Hochreiter and Schmidhuber (1997) presented research using LSTM model, where the research constructed a smart agricultural model and it proved efficient and high accuracy in the prediction of the plant behavior [12]. Ray, S., et al, (2023) presented research named *" An ARIMA-LSTM model for predicting volatile agricultural price series with random forest techniques "*. They introduced hybrid ARIMA-LSTM algorithm to estimate the agricultural crop, using the Random Forest algorithm for choosing adequate inputs lag periods for LSTM. The algorithm obtained good results when performed on three agricultural price series and it proved highly efficient in processing the price oscillation [13].

A related study by Andini & Utomo, (2021) titled *" Climate Prediction Using RNN LSTM to Estimate Agricultural Products Based on Koppen Classification"* aimed to forecast future climate conditions to assess crop yields according to the Köppen classification. Features like temperature, humidity and rainfall were used as inputs in the prediction stage. Using the recurrent neural network, the

ISSN:2302-8149
e-ISSN:2540-9719

research obtained good-accuracy outputs, with a loss value of 0.006 and a Mean Absolute Percentage Errors (MAPE) of 3.29%, [14].

In the area of plant growing prediction**,** Alhnaity, et al. (2021) published a paper named *"An autoencoder wavelet based deep neural networks with attention mechanisms for multi-step prediction of plant growth "*. The researchers supposed an advanced system for predicting plant growing (for example, stem diameter) depending on a time series of IoT sensor. The proposed algorithm used improved LSTM for extracting features, and the obtained outputs stated a significant prediction over traditional models using some standard metrics such as RMSE, MAE, and MAPE [15].

The Naïve Bayes model also proved its high performance in agricultural IoT systems and in enhancing the accuracy of the prediction. Singh, et al. (2023) presented a paper *"IoT and AI-based Intelligent Agricultural Framework for Crop Predictions"***,** supposed an integrated system that links IoT and artificial intelligence (AI) algorithms in order to estimate the most adequate crops depending on the environmental variables. The researchers used a sensor for scaling real-time IoT data in order to determine the range of parameters such as temperatures, humidity, pH levels, and rainfalls, hence measuring the concentration levels of nitrogen, phosphorus, and potassium. The researchers compared many of ML algorithms for determining the best performance in the prediction field like, K Nearest Neighbor, Decision Tree, Naïve Bayes, and Logistic Regression models. The Naïve Bayes model obtained the highest accuracy of 99.39%**,** displaying the highest performance in this field of prediction [16]. Kumar, et al. (2021), designed a platform for predicting soil moisture in the paper *" Smart Agricultures using IoT in Field monitoring and Automations for Soil Moisture using Naive Bayes Prediction "*. The platform based on sensor nodes used to select variables like, climate oscillation and humidity, then applying Naïve Bayes prediction model. The results showed low error rates not exceeding 10%**,** and a strong correlation of up to 90% between estimated and real values, demonstrating the accuracy and reliability of the model in agricultural environments [17].

Table 1 Comparison between the presented research in terms of strengths and research gaps

Study	Field	Technologies Used	Strengths	Research Gaps
Javaid, et al , (2021)	IoB Services	Behavioral analysis	Practical application for customer satisfaction	Does not include agricultural applications
Stary, C. (2020)	IoB Services	Behavioral analysis	Support understand and the dynamic analysis of the organization's performance ecosystem.	Does not include agricultural applications
Ray, et al, (2023)	Agricultural prices	ARIMA + LSTM	High prediction accuracy	Limited to prices, not behavior
Andini, & Utomo, (2021)	Climate and agriculture	RNN-LSTM	Linking climate to production	Ignoring aspects of human behavior
Alhnaity, et al. (2021)	*Agriculture IoT*	IoT + LSTM	Predicting plant growth using real-time IoT data	Ignoring the taxonomic or behavioral aspects of the user-based IoB
Singh, P., et al. (2023)	*Agriculture IoT*	KNN+DT+ Naïve Bayes + LR classifiers.	Predicting suitable crops	Not integrated with user behavior
Kumar, et al. (2021)	*Agriculture IoT*	Naïve Bayes	Predicting soil moisture	Not integrated with user behavior

In general, although most studies have focused on processing Internet of Things (IoT) data from a temporal or categorical perspective, the use of this data within the Internet of Behavior (IoB) framework remains limited. None of these studies have integrated the Internet of Behavior to analyze user behavior with advanced techniques such as LSTM or Naïve Bayes, nor have they addressed prediction in the agricultural sector.

ISSN:2302-8149
e-ISSN:2540-9719

3 Research Method

The proposed methodology consists of a series of structured and integrated steps to predicting and evaluating plant growth behavior using plant data obtained based on Internet of Things (IoT) technologies. The proposed methodology steps (Figure 2) are represented by the following: data preparation, prediction using the LSTM algorithm, classification using Naïve Bayes, and finally, agricultural behavior evaluation.

The first stage involves data preparation and preprocessing. The "Crop Recommendation Datasets" that provided on the Kaggle website. This dataset contains real IoT agricultural data that includes some features like nitrogen percent (N), phosphorus percent (P), and potassium percent (K) in the soil. The determined dataset has been uploaded at program startup, and then 100 random data values have been selected. These values are segmented into training and testing sets (X as input and Y as target data value).

For a good representation of the different features within the agricultural dataset also to keep variety in the presented data and reduce bias in prediction outputs, the stratified random samples procedure was used. The original data is segmented into subgroups (levels) based on common features such as nutrient levels. 100 random samples were selected from each level, divided into 80 real samples for training, 15 for testing, and 5 samples created outside the natural NPK ranges for predicting the randomness of the model.

The next stage is constructing the model based on the LSTM model in order to predict plant growing data values using the following three input features:

1. N (Nitrogen): Soil nitrogen level, ranging from 0 to 140
2. P (Phosphorus): Soil phosphorus level, ranging from 5 to 145
3. K (Potassium): Soil potassium level, ranging from 5 to 205

Based on the outputs of the LSTM model, the prediction value is generated, which is then compared with the optimal reference value. The matching percent (M) between the two values is calculated using the following equations (1) and (2):

$$\beta = (\Gamma - \delta) * \frac{2}{(\Gamma+\delta)} \qquad (1)$$

$$M = 100 - \beta \qquad (2)$$

Where:

Γ :represents the prediction value (the output resulted from the LSTM algorithm),

δ : represents the target value,

β : represents the absolute difference between the two data values,

M : represents the matching percent between the predicted data value and the target value.

Then we move to the next step, which is evaluating the results using Naïve Bayes to classify the growth behavior (suitable or unsuitable). The inputs (N, P, K) will be considered as features (x1, x2, x3), and two classification probabilities are defined:

1. **Probability (suitable)**: When the matching for input features (x1, x2, x3) is greater than or equal to 80% ($M \geq 80\%$), that means the input values (behavior) are suitable. The classification probability for suitability is computed using the Naïve Bayes equation (3):

$$P(s) = \frac{p(x_1/y_1)*p(x_2/y_1)*p(x_3/y_1)*p(y_1)}{p(x_1)*p(x_2)*p(x_3)} \qquad (3)$$

2. **Probability (unsuitable):** When the matching for input features (x1, x2, x3) is less than 80% ($M < 80\%$), that means the input values (behavior) are unsuitable. The classification probability for unsuitability is computed using the Naïve Bayes equation (4):

$$P(u) = \frac{p(x_1/y_2)*p(x_2/y_2)*p(x_3/y_2)*p(y_2)}{p(x_1)*p(x_2)*p(x_3)} \qquad (4)$$

The final classification is chosen based on which of the two probabilities is higher. Also, y_1 is the number of times the matching case repeated for greater than or equal to 80%, y_2 is number of times the

 ISSN:2302-8149
e-ISSN:2540-9719

matching case repeated for less than 80%. Notice that for simplicity, $y_1 + y_2 = 10$ represents the no. of real data that we are selected randomly from Kaggle dataset, and 5 faked data that are considered out of the ranges of the parameters (N P K), so the dataset is partitioned into two parts.

The final step is using the overall results to evaluate whether the input combination (N, P, K) reflects a behavior suitable for plant growth, based on both the LSTM prediction results and the Naïve Bayes classification. This hybrid approach contributes to improving decision-making accuracy within smart agriculture environments.

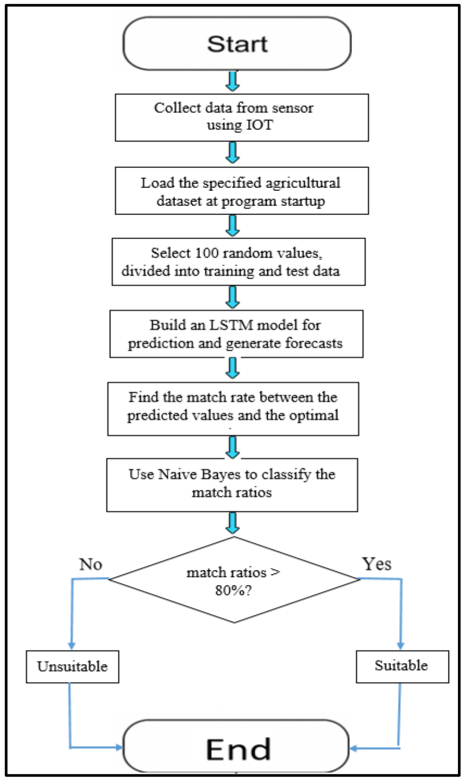

Figure 2 A proposed research workflow diagram

4 Results and Analysis

The sampled dataset was selected randomly from the public dataset and denoted by (H), which means that these values are real data obtained from the Kaggle website, and the output for the plant have to be Healthy as displayed in Table 2:

Table 2 Random input data from IoT dataset

Plant Case	$N(X_1)$	$P(X_2)$	$K(X_3)$
H	23	142	197
	20	141	203
	31	136	197
	17	134	204
	13	121	196
	88	35	40
	81	53	42
	140	38	15
	122	48	16
	118	40	35
	160	4	2

ISSN:2302-8149
e-ISSN:2540-9719

	135	200	0
	150	142	210
U	170	3	205
	160	142	203

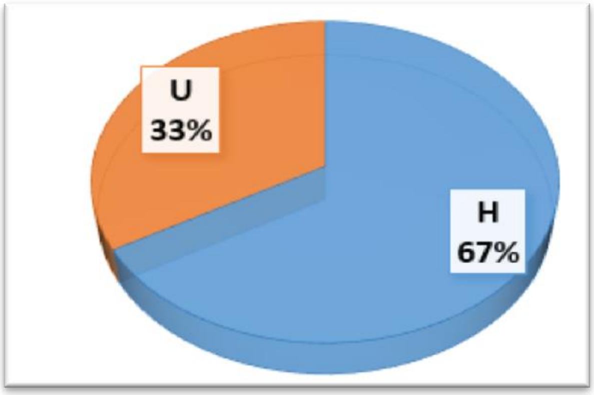

Figure 3 Percentage value for suitable (h) and unsuitable (u) input data

From Figure 3, we noticed that the behavior for the first 10 input data for (N P K) were suitable plant growth (Healthy), where the remaining 5 data were classified unsuitable (Unhealthy) for the plant growth.

Table 3 shows the prediction values generated by the LSTM algorithm and their corresponding optimal values. Figure 4 visualizes the matching percentages.

Table 3 Matching ratio between predictions LSTM sample output and target output data

Predictions (y)	Target Values	(Matching %)
0.8958	0.8628	98.09
0.9164	0.8440	95.82
1.0198	0.9638	85.15
0.8994	0.9588	97.69
0.9779	0.9589	89.33
0.9600	0.9380	91.17
1.1022	0.9918	89.46
1.0376	0.8860	83.43
0.9597	0.8639	91.20
0.9008	0.8483	93.99

ISSN:2302-8149
e-ISSN:2540-9719

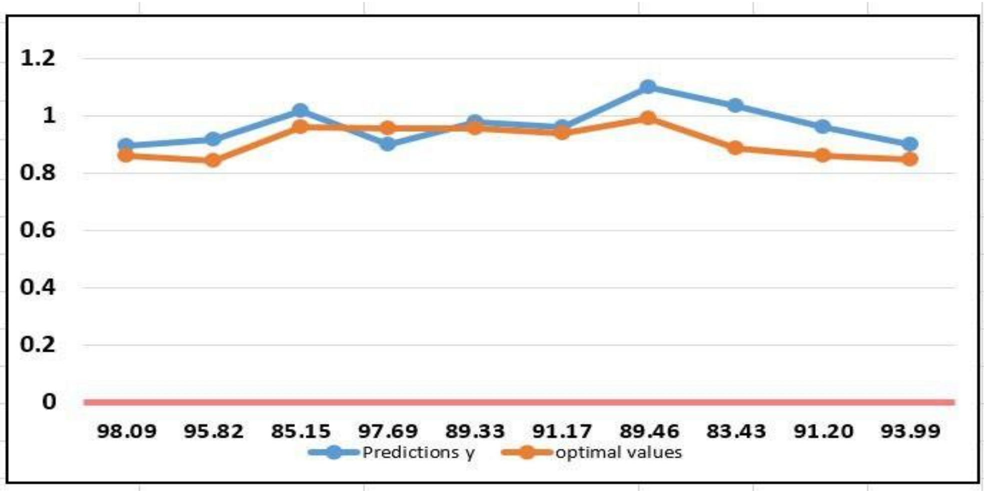

Figure 4 Matching ratio between prediction and optimal (output data)

For increase the accuracy to the proposed algorithm, a Naïve Bayes classifier was applied to the LSTM results. We noticed in Figure 5 and Table 4, the results of all the selected samples from the determined Kaggle dataset indicate that the assessment of the proposed model divides the samples into two regions: first region consists of 10 cases calculated using equation 1 indicate that the classifier correctly identified the behavior as suitable (H) and the second region consists of 5 cases calculated using equation 2 indicate that the classifier correctly identified the behavior as unsuitable (U), so this classification accuracy of the used samples is completely compatible with 100% with the natural behavior of the plant. These values indicate that the proposed hybrid (LSTM, NB) algorithm is suitable and effective as explained in Table 4.

Table 4 Determining the behavior of the plant using naïve bayes classifier (real features)

N (x_1)	***P*** (x_2)	***K*** (x_3)	**Eq.(1) (H)**	**Eq.(2) (U)**	**Assess**
23	142	197	**0.7**	0.3	H
135	200	0	0.24	**0.76**	U
20	141	203	**0.84**	0.16	H
31	136	197	**0.63**	0.37	H
150	142	215	0.12	**0.88**	U
17	134	204	**0.79**	0.31	H
13	121	196	**0.66**	0.34	H
88	35	40	**0.78**	0.22	H
160	4	40	0.45	**0.55**	U
81	53	42	**0.76**	0.24	H
140	38	15	**0.6**	0.4	H
170	3	205	0.28	**0.72**	U
122	48	16	**0.77**	0.23	H
118	40	35	**0.64**	0.36	H
150	142	210	0.15	**0.85**	U

 ISSN:2302-8149
e-ISSN:2540-9719

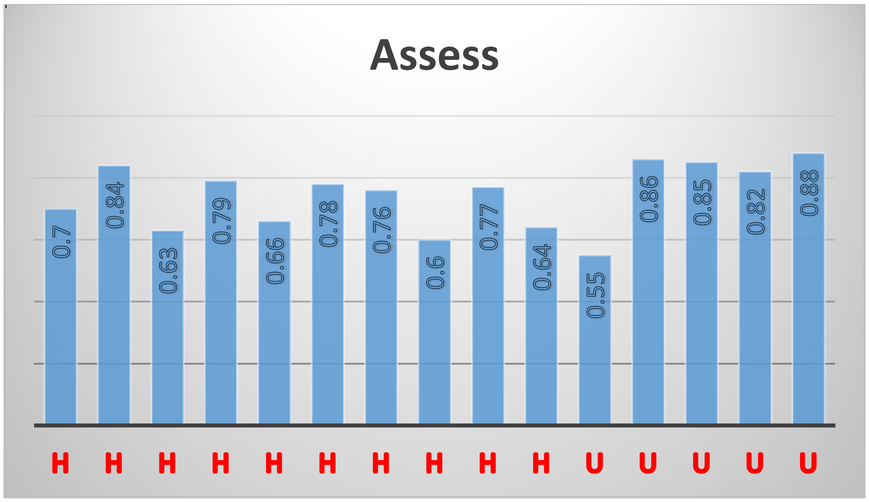

Figure 5 The assessment results using naïve bayes classifier for both equations (1 and 2)

By using the basic prediction metrics: accuracy, precision, recall and F1-score that are used for evaluating the activity of the used ML model and also used for adding more accuracy to the resulted predictions as explained in Table 5. Both models performed well, but the Naïve Bayes classifier slightly go on better than LSTM, this states that the process of combining both models, as done in the proposed hybrid approach, can generally improve the system performance.

Table 5 Prediction metrics values of the proposed ML model

Prediction Metrics	LSTM	NB
Precision (%)	0.88	0.94
Accuracy (%)	0.93	0.96
Recall (%)	0.93	0.96
F1-Score (%)	0.91	0.92

5 Conclusion

In recognition of the importance of agriculture as one of the primary sources of human food and consequently its direct impact on national economies, this research aims to harness artificial intelligence to predict plant growth behavior and productivity. This study proposes a hybrid model using the LSTM network and the Naïve Bayes classifier in the context of the IoB (Internet of behavior) to predict and assess plant growth with environmental data collected from the IoT (Internet of Things) technologies. The main contribution lies in combination of time-series predicting with probabilistic classification for both high accuracy and awareness of agricultural behavior. High consistency of LSTM-based predictions with respect to the optimal values, reaching approximately 91.5%, and robust classification power of the Naïve Bayes. This confirms the activity of the proposed work and highlights its potential as a reliable decision-making tool in smart agriculture applications. However, the study faces some limitations. It relied on a dataset with limited diversity, focusing only on three nutrients (N, P, K). While this contributed to the simplicity of the model structure, it may not fully capture the complexity of real agricultural systems. In addition, external environmental factors such as humidity, temperature, and pH level may affect prediction accuracy under real-world conditions. Future studies may focus on developing the model by incorporating more environmental factors such as temperature, humidity, and pH level, as well as integrating user behavioral patterns. It may also be beneficial to deploy the system in a real smart farming environment to validate its effectiveness in the field and explore opportunities for improvement to make it more adaptive and advanced.

ISSN:2302-8149
e-ISSN:2540-9719

References

[1] Alkhayat, R. Y. Y., Alobaydi, E. K. A., Abdullah, O. M., & Arshad, M. R. H. M. (2022, August). Efficiency and Effectiveness in Utilizing Blended Learning in Context Aware Ubiquitous Learning Systems. In *2022 8th International Conference on Contemporary Information Technology and Mathematics (ICCITM)* (pp. 96-100). IEEE.

[2] Al-kateeb, Z. N., & Abdullah, D. B. (2024). Unlocking the Potential: Synergizing IoT, Cloud Computing, and Big Data for a Bright Future. *Iraqi Journal for Computer Science and Mathematics*, *5*(3), 25.

[3] Javaid, M., Haleem, A., Singh, R. P., Rab, S., & Suman, R. (2021). Internet of Behaviours (IoB) and its Role in Customer Services. *Sensors International*, *2*, 100122.

[4] Oliveira, R. C. D., & Silva, R. D. D. S. E. (2023). Artificial Intelligence in Agriculture: Benefits, Challenges, and Trends. *Applied Sciences*, *13*(13), 7405.

[5] Rather, A. M. (2021). Lstm-based Deep Learning Model for Stock Prediction and Predictive Optimization Model. *EURO Journal on Decision Processes*, *9*, 100001.

[6] Shareef, S. R., & Al-Irhayim, Y. F. U. (2022). Towards Developing Impairments Arabic Speech Dataset using Deep Learning. *Indonesian Journal of Electrical Engineering and Computer Science*, *25*(3), 1400-1405.

[7] Ramani, R., Nimmagadda, P., Rajasekar, S., Unogwu, O. J., Al-Mistarehi, A. H., & Abotaleb, M. (2023). A Novel Long Short-Term Memory (LSTM) Deep Learning IoT Method for Lung Cancer Prediction and Detection. *Full Length Article*, *5*(2), pp. 08-8.

[8] Alobaidy, R., & Al-Talib, G. (2018). Comparative Studying for Opinion Mining and Sentiment Analysis Algorithms and Applications. *Al-Rafidain Journal of Computer Sciences and Mathematics*, *12*(2), pp.13-23.

[9] Bhargavi, P., & Jyothi, S. (2009). Applying Naive Bayes Data Mining Technique for Classification of Agricultural Land Soils. *International journal of computer science and network security*, *9*(8), pp.117-122.

[10] Talib, H. A., Alothman, R. B., & Mohammed, M. S. (2023). Malicious Attacks Modelling: a Prevention Approach for Ad Hoc Network Security. *Indonesian Journal of Electrical Engineering and Computer Science*, *30*(3), pp.1856-1865.

[11] Stary, C. (2020). The Internet-of-Behavior as Organizational Transformation Space with Choreographic Intelligence. In *Subject-Oriented Business Process Management. The Digital Workplace–Nucleus of Transformation: 12th International Conference*, *S-BPM ONE 2020*, *Bremen*, *Germany*, *December 2-3, 2020, Proceedings 12* (pp. 113-132). Springer International Publishing.

[12] Ghojogh, B., & Ghodsi, A. (2023). Recurrent Neural Networks and Long Short-Term Memory Networks: Tutorial and survey. *arXiv preprint arXiv:2304.11461*.

[13] Ray, S., Lama, A., Mishra, P., Biswas, T., Das, S. S., & Gurung, B. (2023). An ARIMA-LSTM Model for Predicting Volatile Agricultural Price Series with Random Forest Technique. *Applied Soft Computing*, *149*, 110939.

[14] Andini, N., & Utomo, W. H. (2021). Climate Prediction using RNN LSTM to Estimate Agricultural Products based on Koppen Classification. *JISA (Jurnal Informatika dan Sains)*, *4*(2), pp.96-99.

[15] Alhnaity, B., Kollias, S., Leontidis, G., Jiang, S., Schamp, B., & Pearson, S. (2021). An Autoencoder Wavelet based Deep Neural Network with Attention Mechanism for Multi-Step Prediction of Plant Growth. *Information Sciences*, *560*, 35-50.

[16] Singh, P., Singh, M. K., Singh, N., & Chakraverti, A. (2023). IoT and AI-based Intelligent Agriculture Framework for Crop Prediction. *International Journal of Sensors Wireless Communications and Control*, *13*(3), 145-154.

[17] Kumar, M. M., Rambabu, P., Srinadh, U., & Satish, T. Smart Agriculture-using IoT in Field Monitoring and Automation for Soil Moisture using Naive Bayes Prediction. *International Journal of Advanced Information Science and Technology (IJAIST) ISSN*, *2319*(2682), pp.20-25.