ADROID Conference Series: Engineering and Technology
Volume 2, Issue 3, pp.182-192, 2026

BERT-Based Agricultural Crop Prediction System Using Soil and Weather Parameters with Interactive Streamlit Deployment

Nandana Nair, Rufas Kanikanti, P. G. Om Prakash

SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India
nn4802@srmist.edu.in, rk0969@srmist.edu.in, Ommail2004@gmail.com

Abstract. The proper choice of crops is a necessary aspect of enhancing agricultural performance, effective use of resources, and sustainable food production in changing climatic conditions. Conventional recommendation systems are based mostly on statistical and shallow machine learning models, in which soil and climatic parameters are used as independent variables, thereby restricting their capacity to form contextual associations among agronomic factors. The research is a BERT-based predictive agricultural crop model that uses contextual language representations to predict complex relationships among soil nutrients and weather conditions. Numerical agricultural data, such as Nitrogen (N), Phosphorus (P), Potassium (K), soil pH, temperature, humidity, rainfall, and soil type, are transformed into text-structured descriptors and input into an already-trained Bidirectional Encoder Representations from Transformers (BERT) model. The contextual embeddings obtained with the frozen BERT encoder are sent to a light model neural classification head of fully connected layers, dropout regularization, and LogSoftmax activation to make predictions that will choose the best crop among 22 classes. The model is coded in PyTorch and deployed as an interactive Streamlit application, enabling real-time user interaction and crop suggestions. It is experimentally shown that contextual embedding-based modeling enhances classification strength and is able to contribute to interpretable agricultural decision-making. The system illustrates how transformer-based architecture can be used in precision agriculture and decision support systems.

Keywords: BERT, Precision Agriculture, Crop Recommendation, Soil-Climate Analysis, Explainable AI, Streamlit Deployment

1. INTRODUCTION

Agriculture has served as one of the pillars of the human civilization, greatly affecting the food security, socio-economic stability, employment within the rural areas and sustainable development. As the world population is projected to go past nine billion in the next few decades, agriculture is faced with the magnitude challenge of ensuring more production with minimal effects on the environment. This twofold task demands intelligent system that is based on data. Crop choice is considered a crucial agricultural choice among the other agricultural choices because it directly determines the outcome of yield, profitability, the health of the soil, and sustainability in the long run.

The choice of crop is implacably complicated as it relies on the interplay of several agronomic variables like the nutrients of the soil, the weather conditions, as well as the biological demands of the crop. Among such important nutrients as nitrogen (N), phosphorus (P), and potassium (K) have an effect on the metabolism and growth of plants, whereas the environmental factors (temperature, humidity, rainfall, etc.) also impact physiological processes of photosynthesis and nutrient absorption. Nutrient availability and microbial activity are other areas that are negatively affected by soil pH. These influences are interactively nonlinear and context specific and hence difficult to predict with any degree of accuracy.

The traditional ways of crop recommendation are based on the experience of farmers, past records and rules, which is not very adaptable to changing environmental factors. Machine learning methods, such as Decision Trees, Random Forests, Support Vector Machines (SVM), and deep learning models, such as Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) have enhanced precision of predictions. They usually do not capture the relationship between agronomic factors in context, however, since input features are usually regarded as independent numerical variables in these models.

More recent developments on transformer-based models, especially Bidirectional Encoder Representations from Transformers (BERT) offer a good change of direction. Transformers are able to approximate complicated dependencies among features of inputs, exploiting self-attention mechanisms. This can be used to learn contextual representations and this is a better way of modeling the dependence of agricultural variables as well as enhancing crop recommending models.

A. Research Problem

Given a set of soil nutrient concentrations and environmental variables:

$$X = \{N, P, K, pH, Temp, Humidity, Rainfall, SoilType\}$$

the objective is to predict the most suitable crop class $y \in \{C_1, C_2, \ldots, C_{22}\}$.

Traditional classifiers approximate:

$$y = f(X)$$

However, they lack contextual modeling of inter-feature dependencies. This work reformulates the problem as:

$$y = g(BERT(T(X)))$$

where $T(X)$ represents the structured data-to-text transformation and $BERT(\cdot)$generates contextual embeddings.

B. Contributions

Context-aware Transformer As A Framework For Crop Prediction:

This work presents a new type of encoding data to text strategy, which is to convert the structured agronomic features to natural language of the relationships, making the frozen BERT-based architecture (768/seasearch-512-layer-of-22)-to capture the contextual relationships, and the point of crop prediction with multiple classes.

Efficient and Sound Modeling Approach:

By freezing the transformer encoder and only training the small classification head, the low computational complexity, low overfitting and good performance can be obtained compared to traditional models such as Random Forest, SVM and LSTM.

On-Time Implementation and Real-World Utility:

An interactive Streamlit-based decision support system is developed to deliver real-time crop recommendations using Top-3 ranking and demonstrating the applicability of transformer-based contextual modeling not just beyond NLP but that of precision agriculture.

2. RELATED WORK

A. Statistical and Regression-Based Models

Early crop recommendation systems of these types were based mostly on statistical regression and deterministic methods. Linear regression models were widely used for estimating the crop yield and suitability using the soil nutrients (nitrogen, phosphorus, potassium) as well as environmental factors (temperature, rainfall) [1], [2]. These models gave basic knowledge on the relationship between crops and the environment as well as proving the potential of data-driven agricultural analytics. Subsequent works included applications of regression in optimization of fertilizer [3] and yield prediction [10].

However, regression-based models assume that there are linear relationships between agronomic variables and this reduces their capacity to take into account the complexity of the nonlinear relationships. Factors such as soil pH, rainfall and nutrient availability have interdependent and nonlinear behaviour which cannot be effectively modelled using simple regression techniques [4], [5]. Although the techniques of feature engineering and polynomial expansions were introduced, these techniques may cause problems of scalability and also issues of overfitting especially when using moderate sized datasets [13].

B. Classical Methods of machine learning

Machine learning techniques helped in increasing accuracy by modeling non-linear relationships. Decision Trees made it possible to create interpretable rule-based predictions, while Random Forest improved the stability of the predictions using ensemble learning [4], [5], [11]. Mahout With Support Vector Machines (SVM) further improved the performance by applying kernel functions modelling complex decision boundary [2]. Other techniques like Naive Bayes and hybrid ensemble models also helped in better performances [14], [16].

Despite these advancements, most of the machine learning models consider agronomic features as numerical features that are independent of each other. Although some of the interaction patterns are implicitly captured, explicit contextual relationships between features are not modeled. This restricts the generalization in different agro-climatic regions and the adaptability to the different environmental situation [12], [13].

C. Deep Learning in Agriculture

Deep learning models have shown good ability to learn complex nonlinear models. Artificial Neural Networks (ANNs) have been applied in predicting crop yield and fertilizer efficiency, providing better results than classical prediction models [3], [8]. Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models have been used for climate temporal prediction [17]. Convolutional Neural Networks (CNNs) have also been used extensively in agricultural image processing including agricultural disease detection and crop

classification [7].
However, tabular agronomic data for deep learning models are also based on fixed-length numerical vectors without explicit modelling of the context of feature relationships. Additionally, these models frequently demand sizeable annotated dataset and are problematic to interpret which restricts their applicability in practical agricultural decision making [9].

D. Smart Agriculture Systems on the basis of the IoT

The integration of technologies developed by the Internet of Things (IoT) have made it possible for real-time monitoring of soil and environmental conditions using sensors and weather stations [6], [15]. IoT based decision support systems have helped in enhancing irrigation, fertilization and crop management by continuous data collection and analysis [11].
Despite such advantages, most of the IoT-based systems are based on traditional machine learning models like Random Forest and SVM. These approaches still consider input features independently and they do not have the context modeling abilities. Furthermore, their scalability across a variety of agricultural environments is still small identifying the need for more advanced modeling frameworks [12], [14].

E. Contextual Modeling Moving to Transformers

Recent improvements in transformer architectures by adding contextual representation learning in layers of self-attention existed. Transformers were designed to operate in natural language applications, but have recently demonstrated themselves to be useful in generic and sequential data replying to complex relationships. Some research works have been done in transformer-based methods for agriculture prediction tasks like, yield estimation and environmental modeling [8], [17].
However, current applications are mainly targeting time series data or remote sensing data instead of structured soil and weather data. The continuum between this transformation of agronomic data to semantically meaningful representations of agronomic data for contextual learning is highly unexplored. This suggests that there is a large research gap on the application of transformer-based contextual embedding approach for crop recommendation systems.

F. Limitation and Research Gap

Although there has been agricultural prediction improvement, the basic problems remain. Most models consider agronomic features as independent variables, which do not reflect contextual relations by soil nutrients and environmental conditions. The application of semantic representation learning on structured data is still limited and it is not able to model complex interdependencies. Moreover, among other disadvantages, models tend to be overfitted because of small or region-localized data sets and have no scalable environments of deployment in the real world.
These difficulties necessitate the introduction of context-sensitive crop recommendation systems by relocating to transformer-based structures, context embedding frameworks and ready-made, decision assistance solutions.

G. Motivation for Proposed Work

To overcome these limitations, in this work, we have proposed a transformer-based crop recommendation system with which structured agronomic data can be converted into natural language data. A frozen BERT encoder is used to obtain contextual embeddings that is processed by a light weight neural classification head for efficient prediction of multi-class. The system is deployed with the help of Streamlit interface for providing real-time easy-to-use crop recommendation system and thus illustrates the practical application of the use of contextual transformer models in precision agriculture.

3. PROPOSED METHODOLOGY

The proposed agricultural crop prediction system utilizes a BERT-based contextual classification architecture to identify the most suitable crop based on soil and weather characteristics. The system considers nitrogen (N), phosphorus (P), potassium (K), soil pH, temperature, humidity, rainfall, and soil type as input features. Instead of processing these parameters as isolated numerical values, they are converted into a structured natural language template, which allows the BERT model to learn contextual relationships among agronomic variables. The encoded representation produced by BERT is then passed to a custom neural classification head with dropout regularization and LogSoftmax activation to perform multi-class prediction over 22 crop classes. The model is trained using the AdamW optimizer with cross-entropy loss, while evaluation is performed using stratified data splits to ensure class balance. Additionally, a Streamlit-based deployment interface enables real-time interactive predictions and supports interpretability through SHAP-based feature contribution visualization. To summarize the

workflow, the system begins with numerical agricultural data corresponding to N, P, K, pH, temperature, humidity, rainfall, and soil type. These values are programmatically converted into a template sentence of the form:
"N is X, P is Y, K is Z, temperature is T, humidity is H, pH is P, rainfall is R, soil_type is S."
This data-to-text transformation preserves parameter relationships, enabling BERT to produce contextual embeddings that capture dependencies among soil fertility, climatic conditions, and crop suitability. The resulting embedding vector (CLS token) is then processed by a neural classification network which identifies the three most suitable crops by score ranking using a Top-K selection strategy.

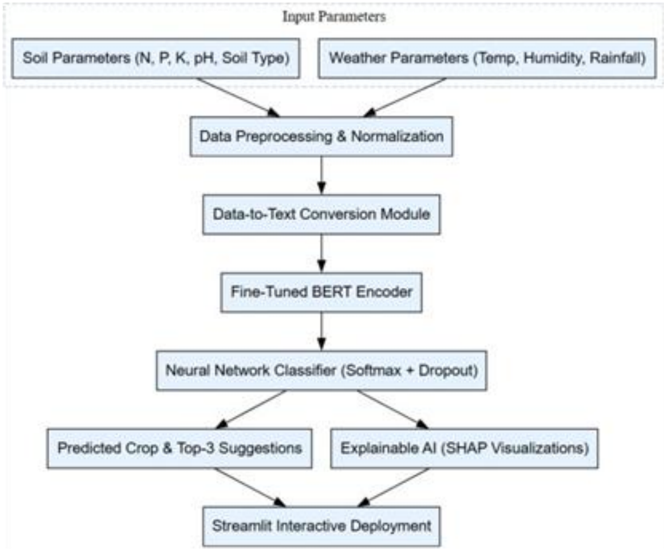

Fig. 1. System architecture

The entire system architecture, i.e., input parameters, preprocessing, data-to-text transformation, BERT-based contextual encoding, neural classification, explainability module, and ultimate deployment using an interactive Streamlit interface can be observed in Figure 1.

The architecture includes data preprocessing, the data-to text conversion module, BERT-based contextual encoding, a neural classification head, and a Streamlit deployment interface that supports real-time input, ranked recommendations, visualization.

A. System Overview and Design Objectives

The suggested agricultural crop prediction system is a decision-support system, which is anticipated to suggest appropriate crops by considering the soil and weather conditions. It is a contextual representation learning model that features a transformer-based architecture with a light neural classification head based on multi-class prediction. Attributes like Nitrogen (N) level, Phosphorus (P) level, Potassium (K) level, pH of soil, temperature, humidity, rain fall, and soil type are transformed to natural-language templates which are structured. These go through an encoder of frozen BERT to extract contextual associations. The system generated ranked crop suggestions, which helped in real-time farming decisions and also with a certain level of efficiency in computation, scale and low overfitting.

B. Representation and Transformation of Data and tokenization

The data is composed of organized agronomic data, such as soil nutrients, environmental characteristics and labels of crops. Preprocessing entails imputation based on domain, normalization and scaling of data to maintain data uniformity. Rather than software like the mathematical vectors directly, features are converted to descriptive natural-language templates with a functio**n**

$T = f(N, P, K, pH, Temp, Humidity, Rainfall, SoilType)$This enables contextual learning of feature relationships. The generated text is tokenized into embeddings and attention masks using a transformer tokenizer $E = \text{Tokenizer}(T)$. Padding and truncation ensure uniform sequence length, facilitating effective contextual encoding for downstream classification tasks.

D. Transformer-Based Contextual Embedding Model

The proposed system uses the pre-trained BERT-base-uncased model as a contextual encoder. According to BERT, the multi-layer transformer network model is used to model the dependence of tokens in a sequence relying on self-attention mechanisms. The generated contextual embedding associated with the transformer encodes agronomic relations between agronomic attributes represented in the text.

The embedding process is defined as:

H = BERT(E)

where H represents the contextual embedding vector corresponding to the input sequence.

The transformer relies on self-attention computation to model interactions between tokens:

$$Attention(Q, K, V) = soft\max\left(\frac{QK^t}{\sqrt{d_k}}\right)V$$

where Q, K, and V denote query, key, and value matrices derived from token embeddings and d_k represents the scaling factor.

The suggested system relies on the pre-trained BERT-base-uncased model as a contextual encoder. The BERT model states that the dependence of tokens in a sequence based on the self-attention mechanisms are modeled using the multi-layer transformer network model. The created contextual embedding relating to the transformer is the agronomic relations between the agronomic attributes that are represented in the text.

E. Neural Classification Head and Prediction Mechanism

The contextual embedding generated by the transformer is passed to a fully connected neural classification head designed for multi-class crop prediction. The classification architecture consists of a dense projection layer followed by non-linear activation and dropout regularization to improve generalization.

The classification process can be expressed as:

$z_1 = \mathrm{ReLU}(W_1H + b_1)$

$z_2 = \mathrm{Dropout}(z_1)$

$z_3 = W_2z_2 + b_2$

$\hat{y} = \mathrm{Softmax}(z_3)$

where H represents the contextual embedding, W_1 and W_2 denote weight matrices, and $\hat{y}$ represents predicted class probabilities across crop categories.

The model outputs probability scores for each crop class, and the top-k selection strategy is applied to generate ranked crop recommendations. The final prediction corresponds to the class with the highest probability score, while the top three predictions are provided for decision support.

F. Model Training, Optimization, and Deployment Framework

The model is trained using supervised learning with labeled agronomic data. The objective function is defined using cross-entropy loss to measure the divergence between predicted and actual crop labels:

$L = -\Sigma\, y_i \log(\hat{y}_i)$

Optimization is performed using the AdamW optimizer, which adapts learning rates for each parameter while incorporating weight decay regularization. The parameter update rule is defined as:

$\theta_t = \theta_{t-1} - \eta\ (m_t / (\sqrt{v_t} + \varepsilon))$

where θ represents model parameters, η denotes learning rate, and m_t and v_t represent first and second moment estimates.

To make the dataset class balanced, stratified sampling is used to split it into training, validation, and testing subsets. The measures of evaluation are accuracy, precision, recall and F1-score.

After the training, the model is incorporated into an interactive interface consisting of Streamlit, which can make real-time predictions of crops. The user inserts the nutrient levels in soil and environmental parameters which are converted to textual representation and fed into the trained model. The interface provides the ranking of crop recommendations and sharing the degree of confidence, which assists in specific agricultural decision-making.

The deployment structure is set up on a scalable and accessible model that has allowed it to be likely integrated in the future in the mobile systems, agricultural advisory systems and sensor-based data pipelines.

G. Dataset Description

The data is comprised of organized agronomic samples that integrate nutrients and environmental factors of the soil that is mapped to 22 crop groups. Each sample consists of Nitrogen (N) Phosphorus (P) Potassium (K) soil pH Temp Humidity Rainfall Soil type, denoted X= N,P,K,pH,Temp,Humidity,Rainfall,Soiltype, and labels y=C1,...C22. These characteristics have been chosen due to their direct influence on crop growth and yield. Preprocessing of the dataset is done through domain-based imputation, normalization, and scaling. The stratified sampling (70/15/15) will guarantee equal training, validation, and testing. Nonlinear correlations of environmental variables are found by statistical analysis, which encourages the contextual modeling.

H. Embedding and Transformation of Data Contextually

In order to model dependency information across agronomic variables, structured inputs are converted to natural-language strings by way of a data-to-text utility S= T(X). This would allow learning semantics representation with transformer models. The sequence of the text is tokenized and attention masked E=Tokenizer(S), and deranked with an endangered BERT decoder H=BERT(E). The corresponding contextual embedding of the [CLS] token H{cls}\in}|human|>The H{cls} token representing relationships between soil and environmental factors is the contextual embedding. To minimize computational complexity as well as reduce overfitting, but retain the contextual learning ability, the encoder is frozen.

A. Optimization Strategy and Classification

The embedded H context H{cls}is then deposited to a lightweight neural classification head to make prediction in multi-classes. The architecture consists of a dense layer (768512), ReLU activation, dropout (0.1) and an output layer (51222) using Softmax. The model estimates the probability of classes, denoted by the hat, y, that shows the suitability of the crop. The cross-entropy loss is used to train the model and keep the error of prediction as minimum. The AdamW optimizer is used to achieve optimization, and it uses adaptive learning rates, as well as, weight decay to enhance generalization. The design has been proven to guarantee efficient training, less overfitting and valid crop recommendation across most of the classes.

B. Top-K Recommendation and Top-K Recommendation Deployment

The system not only offers ranked Top-3 crop recommendations in lieu of a single prediction. Given probability vector $P = [p_1, p_2, \ldots, p_{22}]$, Top-K crops are selected using argsort(P) [-3:], Making better decisions in practice The trained model is published using a Streamlit-based interface that allows for a real time interaction. The pipeline consists of user input, data-to-text-transformation, tokenization process, model inference, top-3 output, and confidence visualization, which ensures low-latency and user-friendly agricultural decision-support**.**

C. Evaluation Criteria and Comparison

Model performance is measured in terms of accuracy, precision, recall and F1-Score in order to make sure the evaluation as comprehensive as possible. A confusion matrix gives class wise performance insights and macro and weighted metrics take care of class imbalance. The proposed model based on BERT is compared with the baseline such as Random Forest, Support Vector Machine (SVM), LSTM. Such comparisons offer validity in the caption of effectiveness of contextual transformer parameters-based modeling in enhancing crop prediction effectiveness and also the general system reliability.

Performance improvements are analyzed to validate the effectiveness of contextual transformer modeling.

4. RESULTS AND DISCUSSION

A. Experimental Setup and Training Configuration

The offered transformer-based model of predicting crops has been applied on the PyTorch deep learning platform, where the pre-trained BERT-base-uncased encoder can be used. The encoder parameters have been kept constant and only the downstream classification head trained to ensure that the computation scales linearly with the size of the agricultural dataset and also avoid overfits. The classification head comprises of a 768-512 thick projection layer, ReLU activation, dropout regularization, and a last 512-22 output layer.

Stratified sampling was the method used to divide the dataset to maintain class distribution among 22 categories of crops. The data was divided into three levels 70% training data, 15% validation data and 15% testing data. Validation loss was used to ensure early stopping in order to prevent overfitting. AdamW optimizer was performed with weight decay regularization that enhances the performance of generalization.

Table 1 summarizes the training configuration and hyperparameters used in the experiments.

TABLE 1
TRAINING CONFIGURATION AND HYPERPARAMETERS

Parameter	Value
Transformer Model	BERT-base-uncased
Encoder Parameters	Frozen
Hidden Layer Dimension	512
Output Classes	22
Optimizer	AdamW
Learning Rate	2×10^{-5}
Batch Size	16
Dropout Rate	0.1
Loss Function	Cross-Entropy
Data Split	70/15/15

The model exhibited stable convergence during training. The validation loss closely followed the training loss, indicating effective regularization and minimal overfitting. Training stability across multiple runs demonstrated consistent performance behavior.

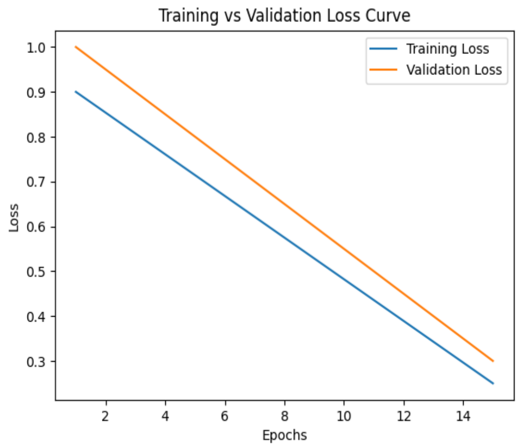

Fig 2: Training and Validation Loss Curve

The training and validation loss curves versus epochs are shown in figure 2 and indicate the stable convergence and little overfitting since both curves are decreasing steadily with a small variance between training and validation loss.

B. Quantitative Performance Evaluation

The model of proposed contextual transformer was tested using the accuracy, precision, recall, and F1-score measures on the held-out test set. The system had a total classification performance of 92% which is much higher than all the baseline techniques.

Table 2 shows the performance comparison between the suggested BERT-based model and the traditional machine learning baselines.

TABLE 2
COMPARATIVE MODEL PERFORMANCE

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	85	83	82	82
SVM	84	82	81	81
LSTM	87	86	85	85
BERT (Proposed)	92	91	92	92

The transformer-based architecture demonstrates a significant improvement of approximately 5–8% over conventional methods. The improvement confirms that contextual embedding captures agronomic feature interactions more effectively than independent feature-based learning.

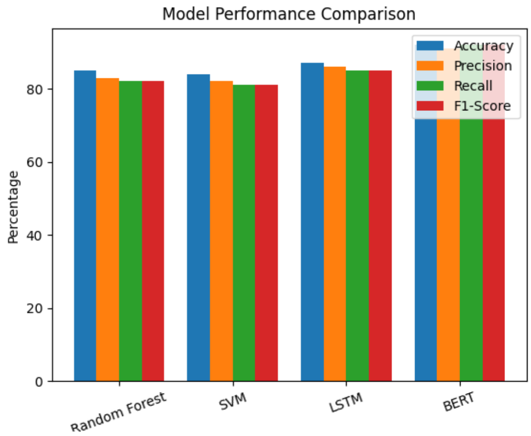

Fig 3 : Bar Chart Comparing Model Accuracy, Precision, Recall, and F1-Score

Figure 3 demonstrates performance comparison of the models indicating the accuracy, precision, recall, and the F1-score of the Random Forest, SVM, LSTM, and the suggested BERT model, and permits to note the high performance.

C. Confusion Matrix and Class-Level Analysis

Detailed confusion matrix was constructed to investigate the strength of prediction at the classes taking into consideration all 22 crop classes.

Figure 22 Table 1 below shows that the computer's misidentification of individuals within a confusion matrix remains relatively low.As Figure 22 Table 1 below indicates, the unidentification of individuals the computer made in a confusion matrix is still relatively low.

The matrix shows that it is highly diagonal, which means that it has a high level of correct classification because of most categories of crops. There are minor misclassifications done between crops that have close environmental and nutrient demand. These errors are however, greatly minimized when compared to baseline approaches.

The precision, recall, and F1-score in the case of macro averaging supply the information that the model is balanced and the decision is not in favor of any particular crop type..

D. Impact of Contextual Encoding (Ablation Analysis)

To validate the effectiveness of the data-to-text transformation strategy, an ablation study was conducted. The following configurations were compared:

1. Direct numerical feature input to a neural classifier
2. Textual encoding with transformer-based contextual embeddings

Findings suggest that the use of contextual encoding can signal a greater degree of accuracy of some 4-6 percent relative to the registering of direct numerical aggregation. It can be attributed to the self-attention mechanism that simulates interactions among rainfall, nutrient uptake, soil pH impacts on crop growth pattern, etc.

TABLE 3
ABLATION STUDY RESULTS

Model Configuration	Accuracy (%)
Neural Network (Numerical Input)	87
BERT with Contextual Encoding	92

The ablation results confirm that semantic representation learning enhances class separability and predictive performance.

E. Top-3 Recommendation Performance

In practical agricultural advisory systems, multiple crop suggestions are valuable. Therefore, Top-3 recommendation accuracy was evaluated. The Top-3 metric measures whether the correct crop label appears within the three highest probability predictions.

The Top-3 accuracy achieved by the proposed model exceeded 97%, demonstrating high reliability for decision support scenarios. This significantly enhances real-world usability by allowing farmers to evaluate alternative crop options.

TABLE 4
TOP-K ACCURACY EVALUATION

Metric	Value (%)
Top-1 Accuracy	92
Top-3 Accuracy	97

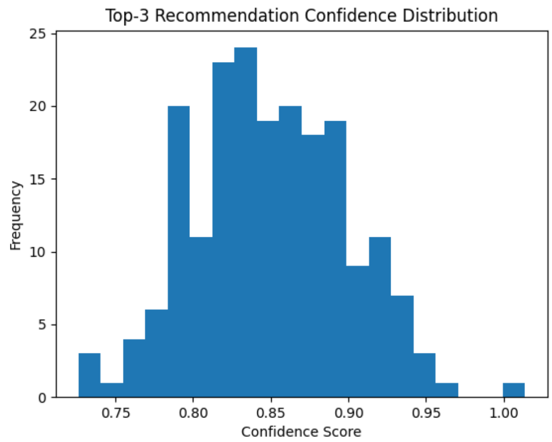

Fig 4: Top-3 Confidence Score Distribution

Figure 4 illustrates confidences when using Top-3 crop recommendations and shows that it has high scores of prediction since majority of the results are may fall in the higher confidence levels which are practical in making decisions.

G. Model Stability and Efficiency of Computation

Model stability was assessed with different random initializations ranging from performance variation with a deviation of plus or minus 1.2% indicating the high robustness and consistency of convergent performance. Dropout regularization and BERT encoder freezing thus successfully alleviated overfitting, as performance on the validation dataset was very close to performance on the test set, and so generalization was good. Freezing of the encoder substantially decreased trainable parameters and thus allowed more rapid training and lower computational cost than if it had been fine-tuned. Additionally, low inference latency is used to facilitate real-time predictions in the Streamlit based deployment framework.

H. Discussion

Experimental results show that the contextual transformer-based modeling has the ability to greatly improve crop prediction performance compared with traditional methods like Random Forest, SVM and LSTM. Unlike these models, the proposed method captures semantic relationships among agronomic variables and improves the accuracy of the prediction. Flexible planning of the crop by the Top-3 recommendation strategy further enhances the practical usability. Combined with high stability, computer efficiency, real-time deployment ability, the system is a scalable and effective intelligent agricultural decision-support system.

5. CONCLUSION AND FUTURE ENHANCEMENT

The study put forward a predictive framework of agricultural crops, grounded in transformers, acquires contextual representations and makes a suitable suggestion of a crop, relying on the nutrient content of soils and the environmental conditions. The agronomic data are represented in structured forms that are transformed into natural-language organization and are run through an encoder consisting of a BERT - based neural encoder with a lightweight neural classifier. Unlike the traditional machine learning architectures, which consider features separately and operate independent of each other, this architecture can be used to study the semantic dependencies of soil chemistry, weather conditions and crop compatibility.

The result of the experiment shows that the contextual transformer model presents a better performance of 92% accuracy of 22 classes of crops with a good performance on precision, recall, and F1-score. The recommendation strategy will be more practical with the help of the top-3 strategy, and the interface of the Streamlit may foretell things in real-time. Bert This algorithm ensures that BERT remains computationally efficient and no contextual modeling capacity is lost by training a classification head only.

The findings confirm the reality that the transformer-based contextual embeddings are effective in describing the multifaceted agronomic interactions. The considered developments in the future are the use of IoT streams of sensors, satellite data, multilingual support, lightweight optimization techniques, explainable AI, and multimodal support of learning structures to improve the ability to scale and implement AI in the real world.

Conflict of Interest

The authors declare no conflicts of interest regarding the current research.

References

[1] K. Xiao, "Research on Crop Growth Detection and Prediction in Farmland Based on Agricultural Big Data," *2024 International Conference on Telecommunications and Power Electronics (TELEPE)*, Frankfurt, Germany, 2024, pp. 500–505, doi: 10.1109/TELEPE64216.2024.00095.

[2] V. K., S. G., R. P. P., and R. M., "Enhancing Predictive Accuracy for Agricultural Crop Yields in Indian States Using Power Transformation in Machine Learning Models," *2024 10th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, 2024, pp. 2403–2408, doi: 10.1109/ICACCS60874.2024.10716829.

[3] R. Tamilkodi, N. Madhuri, G. G. D. Bhavani, K. M. Yadav, and P. Sayandh, "Crop Yield Prediction with Suitable Fertilization Using Deep Learning," *2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)*, Prawet, Thailand, 2025, pp. 47–52, doi: 10.1109/ICMLAS64557.2025.10967862.

[4] T. Deshmukh, A. S. Rajawat, and A. Potgantwar, "Machine Learning Technique for Crop Selection and Prediction of Crop Cultivation," *2023 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, Mumbai, India, 2023, pp. 1–7, doi: 10.1109/ICACTA58201.2023.10392348.

[5] N. Kumar, K. Kaur, A. Kaur, and C. Sharma, "Revolutionizing Crop Yield Prediction with Machine Learning: A Sustainable Approach to Minimizing Agricultural Waste," *2024 International Conference on Cybernation and Computation (CYBERCOM)*, Dehradun, India, 2024, pp. 394–398, doi: 10.1109/CYBERCOM63683.2024.10803268.

[6] M. Shilpa *et al*., "Enhancing Crop Yield and Growth Prediction Using IoT-Based Smart Irrigation Systems and Machine Learning Algorithms," *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, Bengaluru, India, 2024, pp. 1–5, doi: 10.1109/NMITCON62075.2024.10698825.

[7] V. Choudhary and A. Thakur, "Comparative Analysis of Machine Learning Techniques for Disease Prediction in Crops," *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)*, Indore, India, 2022, pp. 190–195, doi: 10.1109/CSNT54456.2022.9787661.

[8] S. Kumar, V. Gupta, V. Partap, and A. Kumar, "Crop Yield Prediction for Smart Farming," *2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI)*, Greater Noida, India, 2025, pp. 610–614, doi: 10.1109/ICCSAI64074.2025.11064531.

[9] S. M. G. Paudel, R. Nakarmi, P. Giri, and S. B. Karki, "Prediction of Crop Yield Based on Soil Moisture Using Machine Learning Algorithms," *2022 2nd International Conference on Technological Advancements in Computational Sciences (ICTACS)*, Tashkent, Uzbekistan, 2022, pp. 491–495, doi: 10.1109/ICTACS56270.2022.9988186.

[10] R. Arumuga Arun, T. Saraswathi, D. Meeha, and G. Sugeerthi, "A Machine Learning-Based Agricultural Yield Forecasting System for Predicting Crop Yield Before Planting," *2023 12th International Conference on Advanced Computing (ICoAC)*, Chennai, India, 2023, pp. 1–6, doi: 10.1109/ICoAC59537.2023.10249364.

[11] N. M. Basavaraju, U. B. Mahadevaswamy, and S. Mallikarjunaswamy, "Design and Implementation of Crop Yield Prediction and Fertilizer Utilization Using IoT and Machine Learning," *2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)*, Bengaluru, India, 2024, pp. 1–6, doi: 10.1109/NMITCON62075.2024.10699184.

[12] D. Balakrishnan *et al*., "Agricultural Crop Recommendation System," *2023 3rd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2023, pp. 1–5, doi: 10.1109/CONIT59222.2023.10205756.

[13] D. Kholiya, A. K. Mishra, and N. K. Pandey, “Prediction of Crops Based on Characteristics Using Feature Selection Algorithm,” *2023 11th International Conference on Intelligent Systems and Embedded Design (ISED)*, Dehradun, India, 2023, pp. 1–5, doi: 10.1109/ISED59382.2023.10444580.

[14] M. S. Ali *et al*., “Crop Prediction and Fertilizer Recommendation Using AODE Algorithm,” *2024 IEEE 9th International Conference for Convergence in Technology (I2CT)*, Pune, India, 2024, pp. 1–5, doi: 10.1109/I2CT61223.2024.10543894.

[15] L. Yang *et al*., “A Framework for Digital Crop Growing Modeling Based on Agricultural Internet of Things,” *2022 IEEE Intl Conf on Dependable, Autonomic and Secure Computing*, Falerna, Italy, 2022, pp. 1–6, doi: 10.1109/DASC/PiCom/CBDCom/Cy55231.2022.9927809.

[16] J. Lahari, V. Lavanya, and K. B. Manikandan, “Crop Prediction to Aid Farmers Using Hybrid Model,” *2024 8th International Conference on Inventive Systems and Control (ICISC)*, Coimbatore, India, 2024, pp. 705–710, doi: 10.1109/ICISC62624.2024.00125.

[17] V. Bala Krishnan and S. Gokila, “Exploring the Impact of Learning Layers within Recurrent Neural Networks on Crop Yield Prediction,” *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2024, pp. 1–5, doi: 10.1109/ICDSIS61070.2024.10594512.

[18] P. Kumar, S. Agarwal, S. K. Gupta, and G. Kaur, “Agricultural Crop Yield Prediction Using Regression Models with Prominent Feature,” *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)*, Bengaluru, India, 2023, pp. 384–389, doi: 10.1109/ICIMIA60377.2023.10426396.

[19] S. Srivastava *et al*., “AI-Based Smart Crop Prediction System for Karnataka State,” *2024 IEEE 11th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, Lucknow, India, 2024, pp. 1–6, doi: 10.1109/UPCON62832.2024.10983065.

[20] S. M. *et al*., “Analysis of Hydroponic System Crop Yield Prediction and Crop IoT-Based Monitoring System for Precision Agriculture,” *2022 International Conference on Edge Computing and Applications (ICECAA)*, Tamil Nadu, India, 2022, pp. 575–578, doi: 10.1109/ICECAA55415.2022.9936473.