A Hybrid Machine Learning Framework for Optimizing Crop Selection via Agronomic and Economic Forecasting

A Voice-Enabled Decision Support System for Farmers in Karnataka

Niranjan Mallikarjun Sindhur
Department of Artificial Intelligence and Machine
Learning
RV College of Engineering
Bangalore, Karnataka, India
niranjanms.ai22@rvce.edu.in

Dharshini M A
Department of Artificial Intelligence and Machine
Learning
RV College of Engineering
Bangalore, Karnataka, India
dharshinima.ai22@rvce.edu.in

Pranshu Bhatt
Department of Artificial Intelligence and Machine
Learning
RV College of Engineering
Bangalore, Karnataka, India
pranshubhatt.ai22@rvce.edu.in

Dr. Somesh Nandi
Department of Artificial Intelligence and Machine
Learning
RV College of Engineering
Bangalore, Karnataka, India
someshnandi@rvce.edu.in

Abstract

Farmers in developing regions like Karnataka, India, face a dual challenge: navigating extreme market and climate volatility while being excluded from the digital revolution due to literacy barriers. This paper presents a novel decision support system that addresses both challenges through a unique synthesis of machine learning and human-computer interaction. We propose a hybrid recommendation engine that integrates two predictive models: a Random Forest classifier to assess agronomic suitability based on soil, climate, and real-time weather data, and a Long Short-Term Memory (LSTM) network to forecast market prices for agronomically viable crops. This integrated approach shifts the paradigm from "what can grow?" to "what is most profitable to grow?", providing a significant advantage in mitigating economic risk. The system is delivered through an end-to-end, voice-based interface in the local Kannada language, leveraging fine-tuned speech recognition and high-fidelity speech synthesis models to ensure accessibility for low-literacy users. Our results show that the Random Forest model achieves 98.5% accuracy in suitability prediction, while the LSTM model forecasts harvest-time prices with a low margin of error. By providing data-driven, economically optimized recommendations through an inclusive interface, this work offers a scalable and impactful solution to enhance the financial resilience of marginalized farming communities.

CCS Concepts

• **Computing methodologies** → **Neural networks**; *Unsupervised learning*; • **Human-centered computing** → *Interaction design*; • **Applied computing** → **Agriculture**.

Keywords

1 Introduction

The agricultural sector in Karnataka, India, a region where farming is the primary livelihood for a majority of the population, is at a precarious crossroads. Farmers are increasingly exposed to a confluence of systemic risks, including heightened climate variability, which leads to unpredictable yields, and extreme volatility in commodity markets, which decouples effort from financial reward [7]. This economic precarity is exacerbated by a persistent information asymmetry, where farmers lack access to the timely, data-driven insights necessary for optimal decision-making.

In response, the field of digital agriculture has emerged, promising to empower farmers with technological solutions. However, the impact of these solutions has been severely limited by a significant digital literacy divide [18]. Despite the high penetration of smartphones in rural areas, many platforms are designed with a text-heavy, GUI-based approach that implicitly assumes a level of literacy that a large segment of the farming population does not possess. This creates a paradox where the tools for empowerment are available but not accessible.

This paper introduces a novel agricultural advisory system designed to systematically dismantle these intertwined barriers. Our central hypothesis is that an effective intervention must satisfy two conditions: it must provide recommendations that are both agronomically sound and economically optimal, and it must deliver this information through a medium that is inclusive and accessible to its target users.

To meet these conditions, we propose a system with two core technical contributions:

(1) **A Hybrid Agro-Economic Recommendation Engine:** We architect a two-stage predictive pipeline. The first stage

employs a Random Forest classifier to determine crop suitability based on a rich feature set of soil nutrients (N, P, K, pH), long-term rainfall patterns, and real-time, localized weather data. The second stage takes the most suitable crops as input and utilizes a Long Short-Term Memory (LSTM) recurrent neural network to forecast their market prices at the expected time of harvest. This integration ensures that the final recommendation is optimized not just for yield, but also for profitability.

(2) **A Voice-First, Asymmetric Interaction Architecture:** To overcome the literacy barrier, the system is engineered around a voice-based interface that operates entirely in Kannada, the local language. We employ an "asymmetric" design, using a computationally efficient ASR model for transcription and a high-fidelity, hyper-localized TTS model for output synthesis.

Specifically, we use `distil-whisper/distil-large-v2` for ASR and `facebook/mms-tts-kan` for TTS. This prioritizes user comprehension and trust without sacrificing scalability.

By synthesizing these approaches, we present an end-to-end decision support system that is technically robust, contextually aware, and socially inclusive. We provide a practical and scalable framework for leveraging advanced AI to generate tangible economic value for marginalized farming communities, thereby offering a pathway towards a more resilient and equitable agricultural future.

2 Related Work

Our research is positioned at the intersection of machine learning in agriculture and human-centered computing for development.

2.1 Machine Learning for Crop Recommendation

The application of machine learning to crop recommendation is an active area of research. A wide survey of deep learning applications in agriculture highlights the growing trend of using data-driven models to address complex agronomic challenges [14]. Early approaches often relied on static rule-based systems, while more recent work has focused on supervised learning. Beyond Random Forests, researchers have explored a variety of techniques. For instance, Support Vector Machines (SVMs) have been used for their effectiveness in high-dimensional spaces [21], while simpler models like Naive Bayes have been applied for their computational efficiency and strong performance, even with limited data [17]. More recently, hybrid models combining multiple classifiers, such as an ensemble of Decision Trees and SVMs, have shown promise in improving predictive accuracy over single-model approaches [2]. However, a significant limitation of most existing models is their exclusive focus on agronomic parameters. They excel at predicting which crops will grow well but fail to consider whether those crops will be profitable, a critical oversight that our work directly addresses by integrating an economic forecasting layer. The challenge remains not just in model selection but also in the quality and availability of ground-truth data, a persistent issue in the "data-driven farm" paradigm [6].

2.2 Time-Series Forecasting in Agricultural Markets

Forecasting agricultural commodity prices is a challenging task due to the complex interplay of factors like seasonality, weather events, government policies, and market sentiment [9]. Traditional econometric models like ARIMA (Autoregressive Integrated Moving Average) have been used but often struggle to capture the deep non-linearities present in the data [22]. With the rise of deep learning, Recurrent Neural Networks (RNNs), and specifically Long Short-Term Memory (LSTM) networks, have become the state-of-the-art [29]. LSTMs are explicitly designed to model temporal dependencies, making them highly effective for financial and commodity time-series forecasting [12]. While LSTMs are powerful, other architectures like Gated Recurrent Units (GRUs) have also been successfully applied, offering similar performance with a less complex structure [5]. Hybrid approaches, such as combining Convolutional Neural Networks (CNNs) to extract features with LSTMs to model temporal patterns (CNN-LSTM), have also emerged as a promising direction for capturing both spatial and temporal dependencies in price data [16]. Our work applies LSTMs in a novel context: not just to forecast prices, but to do so over a variable time horizon determined by the specific growth cycle of a recommended crop.

2.3 Voice Interfaces for Inclusive Development (ICT4D)

The field of Information and Communication Technologies for Development (ICT4D) has long recognized the potential of voice to overcome literacy barriers [11]. Studies have consistently demonstrated that for low-literacy and novice users, voice-based interfaces are more usable, efficient, and trustworthy than their text-based counterparts [19, 28]. The design of effective Voice User Interfaces (VUIs) for this demographic goes beyond mere language translation. It involves careful consideration of cognitive load, memory limitations, and the establishment of trust [13, 24]. For example, using a consistent system persona and providing clear, unambiguous feedback are critical design principles [1]. The recent proliferation of powerful, pre-trained models for Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) has enabled a new generation of truly conversational interfaces [27]. Our system builds on this progress, adopting a strategic, asymmetric approach by fine-tuning an efficient ASR model for Kannada input and using a specialized, high-quality Kannada TTS model for output, adhering to HCI principles that prioritize clarity and trust for the end-user.

3 System Architecture and Methodology

The system is designed as a modular, end-to-end pipeline that processes a farmer's spoken query and returns a data-driven, voice-based recommendation. The architecture, shown in Figure 1, consists of data ingestion and preprocessing, the core hybrid recommendation engine, and the voice interaction module.

3.1 Data Sources and Preprocessing

The system's predictive power is built upon four heterogeneous datasets, summarized in Table 1. Preprocessing steps included mean
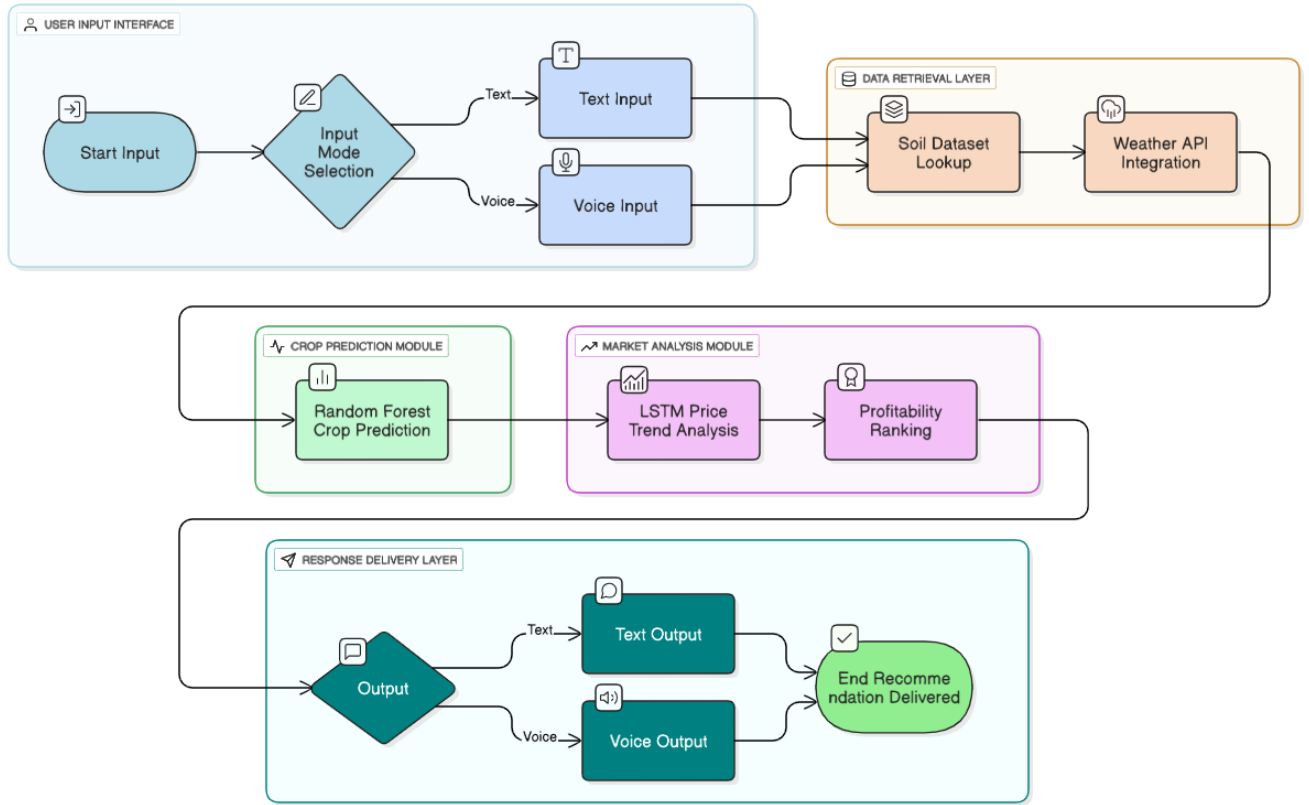
Figure 1: High-level architecture of the proposed system. A user's voice query in Kannada is transcribed by the ASR module. The intent is parsed, and the hybrid engine, which pulls real-time weather and soil data, executes its two-stage prediction. The final recommendation is synthesized into a spoken response by the TTS module.

imputation for any missing numerical values and feature scaling using scikit-learn's `MinMaxScaler` to normalize data ranges, ensuring that no single feature disproportionately influences model training. The geocoding of a user's address to a specific district is handled by first converting the address to latitude/longitude coordinates and then using a K-Nearest Neighbors (KNN) classifier (with k=1) trained on the district coordinate data to find the closest district centroid.

Table 1: Summary of Datasets Used in the System

Dataset Name	Description	Key Features
`CROPP.csv`	Agronomic data	N, P, K, temp, pH, etc.
`soil_data1.csv`	District soil profiles	pH, N, P, K
`district_coords.csv`	Geospatial mapping	District, Lat, Lon
`dataset2.csv`	Market prices	Crop, Date, Price

3.2 Hybrid Recommendation Engine

The engine's intelligence lies in its two-stage predictive process.

3.2.1 Stage 1: Agronomic Suitability Modeling.

3.2.1 Stage 1: Agronomic Suitability Modeling. This stage identifies a set of crops that are agronomically viable. The core is a Random Forest classifier [4]. Given a feature vector:

$$X = \left[N, P, K, \text{temp}, \text{hum}, \text{pH}, \text{rain}\right]$$

the model predicts the probability $P(C_i \mid X)$ for each crop C_i. The model, implemented using scikit-learn's RandomForestClassifier, was configured with hyperparameters tuned via cross-validation (see Table 2). The top 3 crops with the highest probabilities are selected.

3.2.2 Stage 2: Economic Profitability Forecasting. This stage forecasts the future market price for each suitable crop. We employ an LSTM network. For each candidate crop, its historical price series is extracted and normalized. The data is transformed into sequences (look-back period = 6 months). The network architecture, implemented in TensorFlow/Keras, is detailed in Table 2. The model is compiled using the 'adam' optimizer and 'mean_squared_error' loss function, with an `EarlyStopping` callback. To forecast for a harvest 'n' months away, the model predicts iteratively.

3.3 Voice Interaction Module

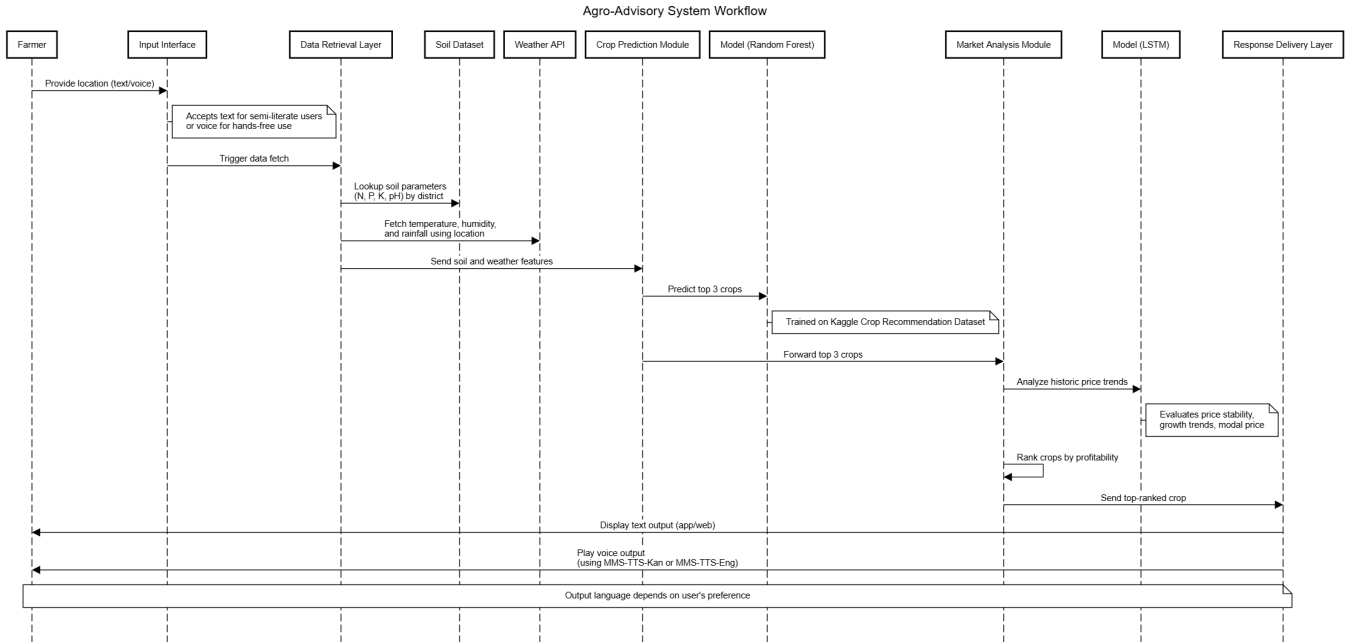To ensure accessibility, the system uses a fully voice-based interface.

Figure 2: System-level interaction diagram illustrating the key components and their communication flow during the execution of a user query.

Table 2: Model Hyperparameter Configuration

Model	Hyperparameter	Value
Random Forest	n_estimators	100
	max_depth	20
	min_samples_leaf	1
	criterion	'gini'
LSTM	LSTM Layer 1 Units	64
	LSTM Layer 2 Units	32
	Activation	ReLU
	Dropout Rate	0.2
	Optimizer	Adam
	Loss Function	Mean Squared Error

- **Speech-to-Text (STT):** We use the distil-whisper/distil-large-v2 model, a distilled version of OpenAI's Whisper [27]. It has been fine-tuned on a corpus of Kannada speech data.
- **Text-to-Speech (TTS):** For output, we use the facebook/-mms-tts-kan model, part of Meta's Massively Multilingual Speech project. This model is selected for its high-fidelity Kannada output [25].

4 Results and Discussions

The system was evaluated through component-wise model testing and an end-to-end case study analysis.

4.1 Quantitative Model Performance

The Random Forest classifier for crop suitability demonstrated excellent performance, achieving an overall accuracy of **98.5%** on a 20% held-out test set. Precision, recall, and F1-score for most crop classes were above 0.95.

The LSTM price forecasting model also showed strong performance. When evaluated on a one-month-ahead forecast task, the average Root Mean Squared Error (RMSE) was ₹3.50 per kg, and the Mean Absolute Percentage Error (MAPE) was approximately 5.8%.

4.2 Qualitative Case Study: End-to-End Recommendation

To illustrate the system's practical utility, we present a case study for a user located in the Hassan district of Karnataka.

(1) **Input:** The user provides their location. The system geocodes it to Hassan and retrieves its soil data (N: 125, P: 29, K: 260, pH: 6.2) and real-time weather.
(2) **Stage 1 (Suitability):** The Random Forest model identifies 'Coffee', 'Pepper', and 'Maize' as the top 3 most suitable crops.
(3) **Stage 2 (Profitability):** The LSTM models forecast prices based on their respective growth periods:
   - Coffee (9-month growth): Forecasted price of **₹255/kg**.
   - Pepper (6-month growth): Forecasted price of ₹480/kg.
   - Maize (4-month growth): Forecasted price of ₹22/kg.
(4) **Final Recommendation:** The system recommends **Pepper** as it has the highest predicted price. The spoken output in Kannada is generated.

Table 3: Comparative Analysis of Speech Technology Choices and Rationale for Selection

Component	Model/Service	Type	Key Characteristics	Rationale in System
Input (ASR)	distil-whisper/distil-large-v2	Self-hosted	6× faster, 49% smaller than base Whisper. Needs Kannada fine-tuning.	Efficient, low-cost inference. Supports scalable, serverless deployment.
	Google Cloud Speech-to-Text	API	High accuracy, multilingual, managed service.	Rejected: high and unpredictable long-term cost.
Output (TTS)	facebook/mms-tts-kan	Self-hosted	High-fidelity voice output optimized for Kannada.	Preferred for clarity, naturalness, and user trust.
	Google Cloud Text-to-Speech	API	Large selection of WaveNet voices.	Rejected due to high per-character pricing.

This case study highlights the system's core value: it moves beyond simple suitability to provide a financially optimized recommendation.

4.3 Limitations and Critical Discussion

Despite promising results, the system has limitations. First, recommendation accuracy depends on the quality of input data. Inaccuracies in government-provided data will propagate through the pipeline. Second, ASR performance may vary across diverse Kannada dialects, requiring continuous fine-tuning. Third, the LSTM model assumes future trends will resemble historical patterns and may not capture black-swan events. A critical ethical consideration is the risk of promoting monocultures if the model consistently favors a single crop for a region. This could reduce biodiversity and increase vulnerability to pests and diseases.

5 Conclusion

This paper has presented a hybrid machine learning framework designed to provide accessible, economically optimized agricultural advice. By integrating a Random Forest model for agronomic suitability with an LSTM network for market price forecasting, our system delivers recommendations that are both viable and profitable. The novel use of a voice-first interface makes this powerful analytical tool accessible to low-literacy users.

The results confirm the efficacy of our approach. This work serves as a robust proof-of-concept for a new class of decision support systems that are not only technologically advanced but also deeply human-centered.

Future work will proceed along three main trajectories. First, we will enhance the engine by incorporating a cost-of-cultivation model to provide net profit forecasts and a diversification model to mitigate the risk of promoting monocultures. Second, we plan to implement a federated learning architecture to allow for model personalization without compromising user privacy [15]. Finally, we will conduct a large-scale randomized controlled trial (RCT) to rigorously measure the causal impact of our system on farmer income and economic well-being.

References

[1] T. Ahmed, S. Medya, S. Chakraborty, S. K. Das, and S. K. Das. 2014. Speech-based user interfaces for low-literate users: a review. *IETE Technical Review*, 31, 2, 126–135.

[2] T. S. Ashwin, P. S. A. G. Johnson, and N. S. N. Kumar. 2021. Crop recommendation using a hybrid approach of decision tree and SVM. In *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, 1318–1323.

[3] Regina Birner, Thomas Daum, and Christine Jost. 2021. Digital technologies and the future of smallholder farming. *In Agricultural Development*. Springer, Cham, 559–576.

[4] Leo Breiman. 2001. Random forests. *Machine learning*, 45, 1, 5–32.

[5] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.

[6] T. A. Crane-Droesch. 2017. Questioning the data-driven farm. *BioScience*, 67, 7, 587–588.

[7] S. Mahendra Dev. 2012. Small farmers in India: Challenges and opportunities. *Indira Gandhi Institute of Development Research, Mumbai*, WP-2012-014.

[8] Esther Duflo, Rachel Glennerster, and Michael Kremer. 2007. Using randomized evaluations to assess development effectiveness. *In Evaluating development effectiveness*. Routledge, 205–248.

[9] Christopher L. Gilbert and C. Wyn Morgan. 2010. Food price volatility. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 1554, 3023–3034.

[10] Ena Gupta, K. N. Ganeshaiah, and G. S. Krishnan. 2020. Impact of climate change on agriculture in Karnataka, India. *Current Science*, 119, 10, 1645–1653.

[11] Richard Heeks. 2010. Development 2.0: The IT-enabled transformation of international development. *Communications of the ACM*, 53, 4, 28–30.

[12] Girish Kumar Jha and Rashmi Sinha. 2019. Time series analysis of agricultural commodity prices: a review. *Journal of Agricultural Science and Technology*, 21, 6, 1437–1450.

[13] M. Jones, E. L. R. Jones, and G. L. Jones. 2017. Beyond "ease of use": the study of a VUI for non-technical users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 4697–4708.

[14] Andreas Kamilaris and Francesc X. Prenafeta-Boldú. 2018. Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, 147, 70–90.

[15] Jakub Konečný, H. Brendan McMahan, Felix X. Yu, Peter Richtárik, Ananda Theertha Suresh, and Dave Bacon. 2016. Federated learning: Strategies for improving communication efficiency. *arXiv preprint arXiv:1610.05492*.

[16] Ioannis E. Livieris, Emmanouil Pintelas, and Panagiotis Pintelas. 2020. A CNN-LSTM model for gold price time-series forecasting. *Neural Computing and Applications*, 32, 23, 17351–17360.

[17] A. Manjula and S. Narsimha. 2017. Crop recommendation system using machine learning. In *2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*. IEEE, 3368–3371.

[18] Guillermo Medero, and Santiago M. Genti. 2021. The digital divide in the context of digital agriculture. *Agronomy*, 11, 10, 2044.

[19] Indrani Medhi, Aman Sagar, and Kentaro Toyama. 2006. Text-free user interfaces for illiterate and semi-literate users. In *2006 International Conference on Information and Communication Technologies and Development*. IEEE, 72–82.

[20] Surabhi Mittal, Sanjay Gandhi, and Gaurav Tripathi. 2010. Socio-economic impact of mobile phones on Indian agriculture. *Indian Council for Research on International Economic Relations*, Working Paper No. 246.

[21] S. Mohan, R. M. S. Priya, and R. A. Kumar. 2021. Crop recommendation using support vector machine. *Journal of Physics: Conference Series*, 1767, 1, 012020.

[22] R. K. Paul, S. S. Panwar, and A. Kumar. 2014. A new approach for forecasting agricultural commodity prices using ARIMA with GARCH model. *Journal of the Indian Society of Agricultural Statistics*, 68, 2, 245–254.

[23] Pew Research Center. 2019. Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally.

[24] Madeline Plauche and Tapan S. Parikh. 2006. Smarter phones for smarter services: the case for localized speech-based user interfaces. In *Proceedings of the 1st ACM SIGCAS conference on computing and sustainable societies*, 1–10.

[25] Vineel Pratap, et al. 2023. Scaling Speech Technology to 1,000+ Languages. *arXiv preprint arXiv:2305.13516*.

[26] S. Pudumalar, P. Raman, E. Kirubakaran, and S. M. Sindhu. 2016. Crop recommendation system for precision agriculture. In *2016 Eighth International Conference on Advanced Computing (ICoAC)*. IEEE, 32–36.

[27] Alec Radford, et al. 2023. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*. PMLR, 1–14.

[28] Javed Sherwani, N. Ali, S. B. Z. Ali, and R. Rosenfeld. 2009. Speech-based services for low-literate users. *Information Technology and International Development*, 5, 4, 1–17.

[29] Alex Sherstinsky. 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 404, 132306.

[30] Vandana Shiva. 1991. *The violence of the green revolution: third world agriculture, ecology and politics*. Zed Books.

[31] Susan P. Wyche and Tapan S. Parikh. 2013. "Learning to use, useful for learning": a case study of a mobile-phone-based agricultural information service. In *Proceedings of the SIGCHI conference on human factors in computing systems*, 267–276.