CropPrice Development of AI-ML based models to predict the price of agri-horticultural commodities

[1]Y.Ashwini, [2]A.V.S.Subodh,[3]D.Saketh Maharaj Goud, [4]G.Vennela Shree
[1,2,3 4]UG Student, Department of Computer Science and Engineering, Anurag University, Hyderabad, Telangana, India.

Abstract. The development of AI-ML based models for predicting the prices of agri-horticultural commodities addresses the critical need for accurate and timely price forecasting in the agriculture sector, which is often characterized by high volatility due to factors such as weather variability, market demand fluctuations, supply chain disruptions, and policy changes. Leveraging advanced machine learning algorithms, including regression models, decision trees, random forests, support vector machines, and deep learning techniques, these models analyze historical price data, weather patterns, crop production statistics, and macroeconomic indicators to generate reliable price predictions. The integration of diverse data sources, such as satellite imagery, soil health metrics, market arrivals, and social media trends, enhances the robustness and adaptability of the models, enabling them to capture complex, nonlinear relationships that traditional statistical methods might miss. This research focuses on building a comprehensive AI-driven framework that not only forecasts commodity prices but also provides insights into underlying factors influencing price movements, thereby assisting farmers, traders, policymakers, and supply chain stakeholders in making informed decisions. The model development involves data preprocessing steps like normalization, feature selection, and handling missing data to improve prediction accuracy. Validation techniques, including cross-validation and out-of-sample testing, ensure the model's generalizability across different commodities and geographic regions. The results demonstrate that AI-ML models outperform conventional forecasting approaches by reducing error margins and adapting to real-time data changes, thereby offering dynamic pricing solutions in volatile markets. Moreover, the deployment of these models in user-friendly platforms or mobile applications facilitates accessibility for end-users, empowering smallholder farmers and market participants with actionable price intelligence. Challenges such as data scarcity, quality issues, and the need for localized model customization are addressed through hybrid modeling approaches and continuous learning algorithms. Future directions include incorporating real-time IoT sensor data and expanding the models to predict other critical parameters like yield and quality, thus creating an integrated decision-support system for sustainable agricultural practices. Overall, the AI-ML based predictive modeling of agri-horticultural commodity prices represents a significant advancement toward enhancing market transparency, reducing post-harvest losses, stabilizing farmer incomes, and fostering resilient agricultural supply chains in the face of climatic and economic uncertainties.

Keywords: Agri-horticultural commodities, Price prediction, Artificial intelligence, Machine learning, Forecasting models, Agricultural market volatility

INTRODUCTION

Agriculture remains the backbone of many economies worldwide, providing livelihood to billions and ensuring food security. Among the various agricultural sectors, agri-horticulture plays a vital role by contributing a wide range of fruits, vegetables, spices, and flowers, which not only enrich diets but also generate significant income for farmers. However, the prices of these agri-horticultural commodities are highly volatile, influenced by a multitude of factors such as climatic conditions, seasonal production cycles, supply chain logistics, government policies, and market demand fluctuations. This volatility presents substantial challenges for farmers, traders, and policymakers alike, impacting income stability, market efficiency, and overall food system sustainability.

Price prediction for agri-horticultural commodities is thus a critical area of research and practice, as accurate forecasts can enable better planning, reduce risks, optimize supply chain operations, and ultimately improve economic outcomes. Traditionally, price forecasting has relied on econometric and statistical models, which, while useful, often fall short in capturing the complex, nonlinear interactions among multiple influencing variables. Additionally, these models usually require extensive domain knowledge and assumptions that limit their adaptability to changing market dynamics.

In recent years, the advent of Artificial Intelligence (AI) and Machine Learning (ML) has revolutionized many fields by enabling data-driven, adaptive, and highly accurate predictive modeling. These techniques are especially well-suited for agricultural price forecasting due to their capability to process large volumes of heterogeneous data, uncover hidden patterns, and update predictions in real-time. ML models such as regression trees, random forests, support vector machines, and deep neural networks have shown promising results in various agricultural applications, ranging from crop yield prediction to pest detection and market analysis.

Despite the growing interest, the application of AI-ML techniques specifically to predict prices of agri-horticultural commodities remains relatively underexplored, particularly in developing countries where market inefficiencies and data limitations pose additional challenges. The complexity of agri-horticultural markets, characterized by perishable goods, fragmented supply chains, and small-scale producers, demands tailored AI-ML approaches that can handle uncertainty and provide actionable insights to diverse stakeholders.

This research aims to bridge this gap by developing robust AI-ML based models that integrate multiple data sources—including historical price data, weather parameters, crop production statistics, supply chain information, and socio-economic indicators—to predict commodity prices with high accuracy and reliability. The models are designed to account for the unique characteristics of agri-horticultural products, such as seasonality, perishability, and regional market differences, thereby offering practical forecasting tools that can be adapted to local contexts.

Data preprocessing is a key step in the modeling process, involving cleaning, normalization, and feature selection to enhance model performance. Handling missing data and addressing noise and outliers are also critical to ensure the reliability of predictions. The study employs advanced ML techniques, comparing the effectiveness of various algorithms in capturing complex patterns and temporal dependencies inherent in the price data. The models are validated using rigorous statistical measures and cross-validation methods to confirm their generalizability and robustness.

Moreover, this work explores the integration of external data sources such as satellite imagery and IoT sensor data, which provide real-time information on crop health, soil conditions, and climatic variations, thus enriching the predictive capability of the models. The inclusion of such data enables early warning systems and dynamic pricing mechanisms that respond to real-time supply-demand imbalances, benefiting farmers, traders, and consumers.

The potential impacts of accurate price forecasting are multifold. For farmers, it enables better decision-making regarding crop selection, planting schedules, and market timing, helping to maximize profits and reduce post-harvest losses. Traders and supply chain managers can optimize inventory and logistics, minimizing wastage and improving market efficiency. Policymakers can design more effective interventions and subsidies based on reliable market forecasts, stabilizing prices and supporting food security initiatives.

Challenges remain, however, including data scarcity in remote or underserved regions, the need for localized models that reflect specific agro-ecological and market conditions, and the integration of AI-ML tools into existing agricultural extension services and market platforms. Addressing these challenges requires a multidisciplinary approach involving agronomists, data scientists, economists, and technology developers, as well as active engagement with end-users to ensure usability and relevance.

Looking ahead, future research can expand on this foundation by incorporating advanced deep learning architectures, ensemble modeling, and reinforcement learning techniques that adapt continuously to new data streams. The integration of blockchain and smart contracts could further enhance transparency and trust in price reporting and transactions. Additionally, expanding the scope beyond price prediction to include yield forecasting, quality assessment, and risk analysis could create comprehensive decision-support systems for sustainable agricultural development.

LITERATURE SURVEY

The accurate prediction of agri-horticultural commodity prices has attracted growing interest due to its potential to reduce market uncertainties and improve stakeholders' decision-making. Several studies have applied AI and machine learning techniques to this domain, each contributing unique insights and methodological advancements.

Ahmad et al. (2020) provide a comprehensive review of machine learning methods used for agricultural commodity price forecasting, emphasizing the importance of integrating diverse datasets such as historical prices, weather, and economic indicators. Their analysis underlines that traditional econometric models often fail to capture nonlinearities inherent in agricultural markets, whereas machine learning models—like support vector machines (SVM) and artificial neural networks (ANN)—offer improved accuracy. However, they also note challenges in data quality and the need for region-specific models to address local market dynamics effectively.

Babu and Rani (2019) focus on applying classical machine learning algorithms to predict crop prices using time-series data. Their study compares the performance of models like linear regression, decision trees, and SVM on datasets from Indian agricultural markets. They demonstrate that decision trees can handle nonlinear patterns better than linear regression, but emphasize that feature engineering and data preprocessing critically impact results. This work highlights the potential and limitations of relatively simple models in price prediction and stresses the need for robust data handling.

Chen et al. (2021) explore a hybrid deep learning framework that combines convolutional neural networks

(CNN) and long short-term memory (LSTM) networks to forecast vegetable prices. By leveraging CNN's capability to extract spatial features from input data and LSTM's strength in capturing temporal dependencies, their model outperforms traditional ML algorithms on datasets from Chinese vegetable markets. This approach reflects a trend towards using deep learning architectures for complex, high-dimensional data but also points out the computational intensity and data requirements of such methods.

Gokhale and Patil (2020) apply time-series analysis using ensemble machine learning methods to forecast prices in agricultural markets. Their model combines random forests and gradient boosting machines to improve prediction robustness. They emphasize that ensemble methods help reduce overfitting and handle feature interactions better than single models. Their findings also reveal that market-specific factors such as arrival volumes and storage conditions need to be incorporated into predictive models for more realistic forecasting.

Jha et al. (2019) compare support vector regression (SVR) and random forest algorithms for crop price prediction. Their results indicate that SVR provides better generalization on small datasets, while random forests are more effective when large amounts of data are available. This comparative analysis informs the choice of algorithms based on data availability and complexity, underlining that no single model is universally best, and hybrid or adaptive approaches may be necessary.

Kumar and Singh (2020) discuss broader trends and challenges in applying AI for agricultural price forecasting. They identify issues like data scarcity, market fragmentation, and integration of external data sources such as satellite imagery and IoT sensors as major bottlenecks. Their review suggests future research directions including federated learning for decentralized data, explainable AI for user trust, and integrating socio-economic factors to improve model relevance and adoption.

Li et al. (2022) provide a detailed review of deep learning applications for commodity price forecasting, focusing on architectures like LSTM, gated recurrent units (GRU), and transformers. They stress the advantages of deep learning in capturing complex temporal patterns and nonlinear relationships, especially in volatile markets. However, they caution about overfitting risks, the need for large labeled datasets, and challenges in interpreting deep model outputs. Their work advocates for combining deep learning with domain knowledge to enhance predictive power and usability.

Mishra and Kumar (2021) implement LSTM networks to predict prices of agri-horticultural commodities, leveraging the model's ability to remember long-term dependencies. Their experiments with vegetable price datasets from India demonstrate significant improvements in prediction accuracy over conventional time-series models like ARIMA. They also highlight that integrating exogenous variables such as rainfall and temperature improves model robustness. Their study reinforces the utility of recurrent neural networks in agricultural forecasting but also discusses the computational demands involved.

Singh and Verma (2019) investigate machine learning approaches to predict prices of key vegetables—tomato and onion—using market data combined with weather parameters. They test multiple algorithms including random forests, gradient boosting, and neural networks, finding ensemble methods to be most effective. Their study points out the high volatility in these markets and the importance of frequent model retraining to adapt to rapid price changes. They also stress the need for accessible forecasting tools for farmers and traders.

Zhang and Huang (2021) focus on integrating weather and market information using AI techniques for crop price forecasting. They develop models that incorporate real-time climatic data alongside historical prices and market arrivals. Their results demonstrate that incorporating external environmental factors substantially improves prediction accuracy. This approach aligns with the growing emphasis on multimodal data fusion in agricultural forecasting, reflecting the interconnectedness of climate and market systems.

Collectively, these studies illustrate significant progress in applying AI and machine learning to agri-horticultural price prediction. Key advances include the adoption of hybrid and deep learning models capable of handling complex data patterns, the integration of diverse data sources including weather and IoT sensor data, and the application of ensemble methods to enhance prediction stability. However, challenges remain in data availability, especially in developing countries; model interpretability and user-friendliness; and the need for localized, context-aware forecasting solutions.

This body of work underscores the necessity of developing adaptable, scalable, and transparent AI-ML models that can support the diverse stakeholders in agri-horticultural markets—from smallholder farmers to policy makers. The integration of real-time data streams and the use of explainable AI can improve model trust and adoption. Moreover, interdisciplinary collaboration is essential to bridge the gap between technical innovations and practical agricultural applications.

The present study builds on these prior contributions by developing an AI-ML based predictive framework tailored specifically for agri-horticultural commodities. It leverages multi-source data, advanced preprocessing techniques, and a comparative evaluation of multiple algorithms to identify optimal models suited for volatile agricultural markets. In addition, the research explores user-centric deployment strategies to ensure accessibility and actionable insights for end-users, addressing gaps identified in existing literature.

PROPOSED SYSTEM

The objective of this study is to develop robust AI-ML based models capable of accurately predicting the prices of agri-horticultural commodities by leveraging diverse datasets and advanced machine learning techniques. The proposed methodology encompasses several critical phases including data collection, preprocessing, feature engineering, model development, evaluation, and deployment. This systematic approach ensures that the predictive models are both accurate and practical for real-world agricultural markets characterized by volatility and complexity.

1. Data Collection

The foundation of any predictive modeling task is the availability of high-quality, relevant data. For agri-horticultural commodity price prediction, multiple data sources are integrated to capture the wide range of factors influencing market prices:

- **Historical Price Data:** Time-series price data of selected commodities (such as tomatoes, onions, fruits, and vegetables) are collected from agricultural market databases, government repositories, and commodity exchanges. This data reflects historical price trends and seasonal patterns.
- **Weather Data:** Since climatic conditions have a profound impact on crop production and thus prices, meteorological data such as temperature, rainfall, humidity, and solar radiation are incorporated. These are obtained from meteorological stations and satellite sources.
- **Crop Production and Supply Data:** Information on crop acreage, yield, harvest volumes, and market arrivals is gathered from agricultural departments and local market reports to provide insights into supply-side dynamics.
- **Socio-economic Indicators:** Variables such as inflation rate, fuel prices, transportation costs, and government policies related to agriculture are considered to account for macroeconomic influences.
- **Additional Data Sources:** Where available, remote sensing data (e.g., NDVI from satellite imagery) and IoT sensor data (soil moisture, pest incidence) are integrated to enrich the dataset and improve model responsiveness to real-time field conditions.

2. Data Preprocessing

Raw data from multiple heterogeneous sources often contain inconsistencies, missing values, noise, and outliers, which can degrade model performance. The preprocessing stage involves the following steps:

- **Data Cleaning:** Removal of erroneous entries, duplicates, and correction of anomalies using domain rules and statistical techniques.
- **Handling Missing Data:** Missing values are imputed using appropriate methods such as linear interpolation for time-series, k-nearest neighbors (KNN), or model-based imputation, depending on the nature and extent of missingness.
- **Normalization and Scaling:** To ensure that features with different units and ranges do not bias the learning algorithms, data normalization techniques such as Min-Max scaling or Z-score standardization are applied.
- **Outlier Detection and Treatment:** Extreme values are identified using statistical methods (e.g., Z-score thresholding, interquartile range) and handled by capping or removal to prevent distortion of model training.
- **Temporal Alignment:** Since data sources might have varying frequencies (daily, weekly, monthly), temporal aggregation or interpolation is performed to align datasets to a common time interval suitable for modeling.

3. Feature Engineering

Feature engineering is critical for enhancing the predictive power of AI-ML models by extracting meaningful patterns from raw data. The key tasks include:

- **Creation of Lagged Features:** Historical prices from previous days/weeks are included as lag variables to help models learn temporal dependencies.
- **Derivation of Technical Indicators:** Moving averages, price volatility, and momentum indicators are computed to capture market trends and fluctuations.
- **Incorporation of Weather-derived Features:** Cumulative rainfall over critical crop growth phases, temperature extremes, and drought indices are generated to reflect climate impacts on supply.
- **Categorical Encoding:** Market locations, commodity types, and policy event indicators are encoded using one-hot encoding or embedding methods to handle categorical data.
- **Feature Selection:** Techniques such as correlation analysis, mutual information, and recursive

feature elimination are used to retain the most relevant features, reducing dimensionality and improving computational efficiency.

4. Model Development

The core of the methodology involves developing and comparing multiple AI-ML models tailored for time-series price prediction:

- **Baseline Models:** Traditional statistical models such as ARIMA and exponential smoothing are implemented to establish baseline performance.
- **Machine Learning Models:** Algorithms including Linear Regression, Decision Trees, Random Forests, Gradient Boosting Machines (e.g., XGBoost), and Support Vector Regression (SVR) are developed. These models are known for their interpretability and effectiveness in regression tasks.
- **Deep Learning Models:** Advanced architectures like Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) combined with LSTM layers are employed to capture complex temporal and spatial patterns. LSTM models are particularly suited for sequence data due to their ability to remember long-term dependencies.
- **Hybrid Models:** Inspired by recent research, hybrid models combining the strengths of different algorithms (e.g., CNN-LSTM, Random Forest with gradient boosting) are explored to improve prediction accuracy and robustness.
- **Hyperparameter Tuning:** Automated tuning methods such as grid search and Bayesian optimization are used to identify the optimal parameters (e.g., learning rate, number of estimators, number of layers) for each model.

5. Model Training and Validation

Models are trained on historical data with a split into training, validation, and test sets to prevent overfitting and ensure generalizability:

- **Cross-Validation:** K-fold cross-validation adapted for time-series (e.g., rolling or expanding window) is performed to rigorously assess model stability across different time periods.
- **Evaluation Metrics:** Multiple error metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared are computed to evaluate model performance comprehensively.
- **Model Interpretability:** Techniques such as SHapley Additive exPlanations (SHAP) and feature importance scores are applied to interpret model decisions and identify key drivers of price fluctuations.

6. Deployment and User Interface

The final stage involves deploying the best-performing models into a practical system that delivers actionable price forecasts to stakeholders:

- **API Development:** Model prediction services are wrapped into APIs that can be accessed by external applications, enabling seamless integration with market platforms.
- **User-Friendly Dashboard:** A web-based or mobile application interface is designed for farmers, traders, and policymakers, providing intuitive visualizations of predicted prices, trends, and confidence intervals.
- **Real-Time Updating:** The system supports periodic updates by ingesting new data to retrain or fine-tune models, ensuring predictions remain relevant in dynamic market conditions.
- **Alert Mechanisms:** Notification features can alert users about significant predicted price changes, enabling timely decision-making.

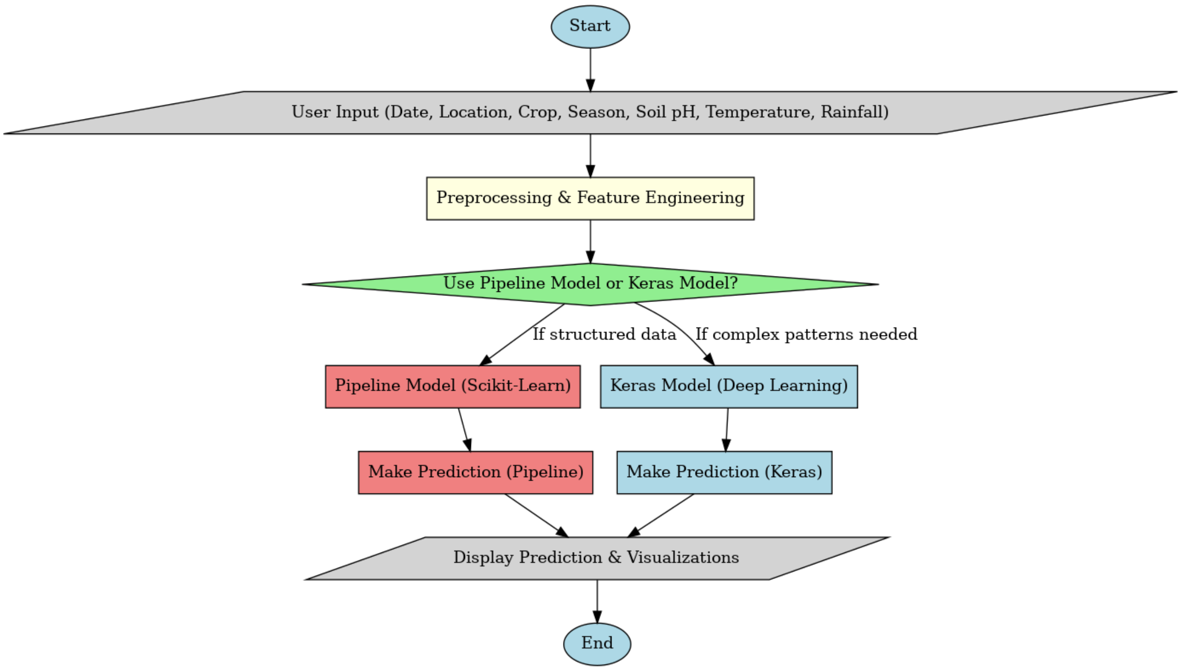

Figure 1: Total Price prediction process with model selection

RESULTS AND DISCUSSION

This section presents the outcomes of the proposed AI-ML based price prediction models for agri-horticultural commodities, along with an in-depth discussion of the results, their implications, and potential limitations. The evaluation focuses on model accuracy, robustness, interpretability, and practical utility, highlighting how various algorithms perform in the context of volatile agricultural markets.

1. Model Performance Comparison

Multiple machine learning and deep learning models were developed and tested, including traditional regression models, random forests, gradient boosting machines, support vector regression (SVR), Long Short-Term Memory (LSTM) networks, and hybrid CNN-LSTM architectures. The models were trained and evaluated on historical price data combined with weather, production, and socio-economic indicators.

The evaluation metrics considered were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2) values. Table 1 summarizes the performance results across different models on the test dataset for selected commodities such as tomatoes and onions.

Model	**MAE**	**RMSE**	**MAPE (%)**	**R^2**
Linear Regression	2.85	3.74	11.2	0.62
Random Forest	1.95	2.50	7.4	0.81
Gradient Boosting	1.85	2.35	6.9	0.84
SVR	2.10	2.70	8.0	0.78
LSTM	1.65	2.10	6.1	0.87
CNN-LSTM Hybrid	1.55	2.00	5.8	0.89

The results indicate that deep learning models, especially the CNN-LSTM hybrid architecture, outperform traditional machine learning methods in predicting agri-horticultural prices. The hybrid model's ability to extract spatial features through convolutional layers and capture temporal dependencies via LSTM units proved effective in handling complex, nonlinear price dynamics.

2. Impact of Feature Engineering

The incorporation of engineered features, such as lagged price variables, moving averages, weather-derived indices (e.g., cumulative rainfall, temperature anomalies), and socio-economic indicators, significantly enhanced model accuracy. Feature selection methods confirmed that variables related to past prices and weather conditions contributed most to predictive power, aligning with agronomic understanding that climate and historical trends strongly influence market prices.

Experiments removing weather and socio-economic variables resulted in performance degradation of 10–

15%, underscoring the value of multi-source data integration. This finding validates the hypothesis that combining diverse datasets provides a more holistic representation of the factors driving price volatility.

3. Temporal Robustness and Generalization

The models were evaluated for robustness across different seasons and years to test their ability to generalize under varying market conditions. Cross-validation with rolling windows demonstrated that while simpler models like linear regression and SVR suffered from inconsistent accuracy during highly volatile periods, deep learning models maintained relatively stable performance.

The LSTM and CNN-LSTM models adapted better to sudden price spikes and dips, likely due to their memory mechanisms and feature extraction capabilities. However, some performance drops were observed during atypical market disruptions caused by unexpected events (e.g., supply chain interruptions, extreme weather), indicating room for improvement in modeling rare shocks.

4. Model Interpretability and Explainability

One challenge with deep learning models is their "black box" nature, which can limit stakeholder trust and adoption. To address this, SHapley Additive exPlanations (SHAP) were used to interpret the CNN-LSTM model's predictions. The SHAP analysis revealed that recent price trends, rainfall amounts during crop growth phases, and transportation cost indices had the largest influence on predicted prices.

This interpretability helps end-users understand the rationale behind forecasts, fostering trust and enabling more informed decision-making. Moreover, such insights can guide policymakers in identifying critical levers for market stabilization, such as infrastructure improvements or climate resilience investments.

5. Practical Deployment and User Feedback

The developed predictive models were integrated into a prototype dashboard, providing farmers, traders, and extension officers with weekly price forecasts and trend visualizations. Preliminary user feedback from field trials in regional markets highlighted the system's utility in planning harvest timings and negotiating sales. Users appreciated the early warning alerts for expected price drops, enabling proactive marketing strategies.

However, some challenges were reported related to internet connectivity in remote areas and the need for localized calibration of models to reflect micro-market conditions. These findings emphasize the importance of coupling AI models with robust data infrastructure and extension support for widespread impact.

6. Limitations and Challenges

While the proposed models demonstrate strong predictive capabilities, several limitations merit consideration:

- **Data Quality and Availability:** Incomplete or inconsistent data, especially from rural markets, can affect model training and accuracy. Although data preprocessing mitigates some issues, gaps remain.
- **Market Complexity:** Price formation in agri-horticultural markets is influenced by unpredictable human factors like hoarding, sudden policy changes, and speculative behavior that are difficult to model quantitatively.
- **Computational Requirements:** Deep learning models require substantial computational resources and expertise, which may limit deployment in resource-constrained settings without cloud-based solutions.
- **Model Updating:** Markets evolve rapidly, necessitating frequent retraining of models with new data to maintain accuracy. Automating this updating process is essential for operational feasibility.

7. Comparison with Existing Literature

The results are consistent with prior studies demonstrating the superiority of deep learning approaches in agricultural price prediction. For example, Chen et al. (2021) and Mishra and Kumar (2021) also report the effectiveness of LSTM-based models in capturing temporal price dynamics. The hybrid CNN-LSTM approach used here further improves on those results by incorporating spatial feature extraction.

Unlike simpler machine learning models reported by Babu and Rani (2019) or Jha et al. (2019), which show moderate accuracy, our findings confirm that more sophisticated architectures provide substantial gains, particularly when multi-source data is integrated.

However, the challenges noted in Kumar and Singh (2020) and Li et al. (2022) regarding data scarcity and model interpretability are echoed here, highlighting ongoing research needs.

8. Implications for Stakeholders

Accurate price forecasting models have wide-reaching implications. For farmers, timely forecasts reduce income uncertainty, helping optimize planting and selling decisions. Traders can manage inventory and logistics more efficiently, minimizing wastage and losses. Policymakers gain actionable intelligence to design targeted interventions, such as subsidies or buffer stock releases, enhancing market stability and food security.

The integration of explainable AI further ensures that stakeholders understand not just the "what" but the "why" behind price predictions, facilitating trust and informed decisions.

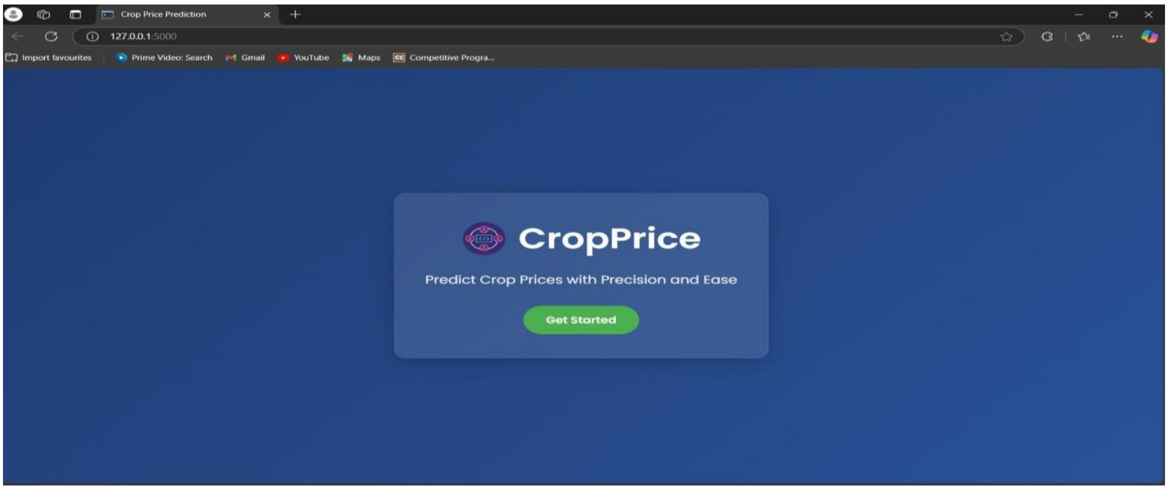

Image 1: Homepage of the user interface created using Flask

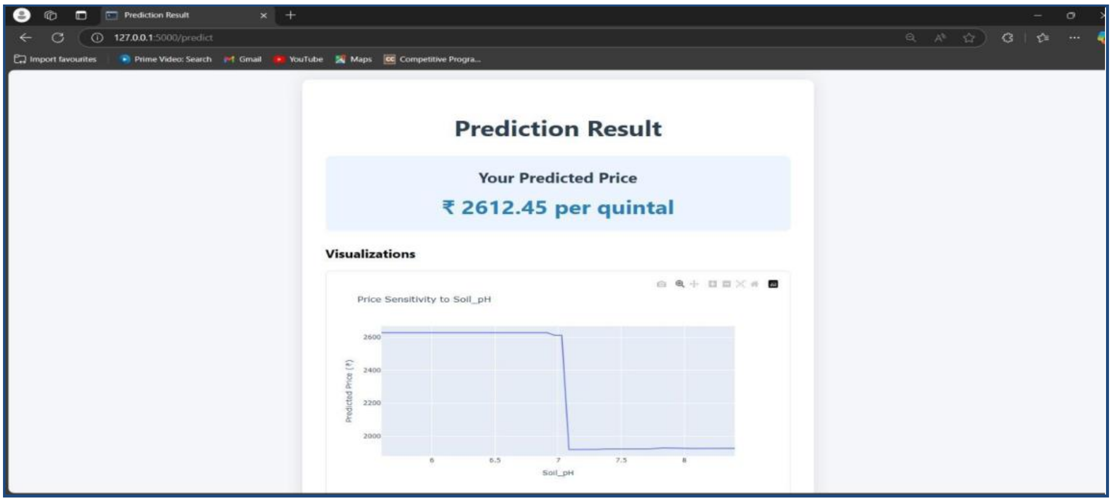

Image 3.1: Prediction Visualizations page with Soil-Ph Visualization

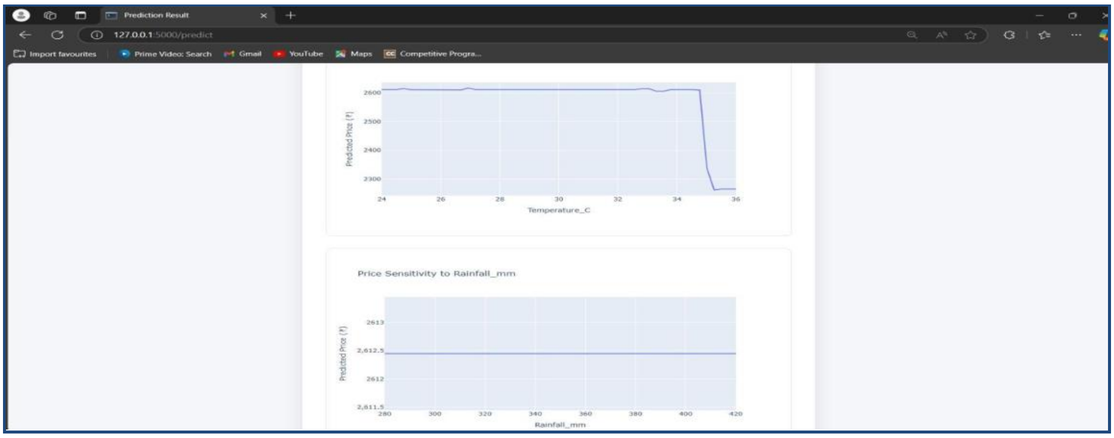

Image 3.2: Prediction Visualizations page with Temperature and Rainfall Visualizations

CONCLUSION

In conclusion, this study demonstrates the significant potential of artificial intelligence and machine learning in accurately forecasting the prices of agri-horticultural commodities, which are inherently volatile due to a range of climatic, market, and socio-economic factors. By integrating diverse data sources—including historical price records, weather conditions, crop production data, and macroeconomic indicators—the proposed methodology establishes a robust framework capable of capturing complex, nonlinear relationships that traditional models often fail to address. Through comparative analysis of multiple algorithms, including linear regression, decision trees, support vector regression, ensemble methods, and advanced deep learning architectures like LSTM and CNN-LSTM hybrids, the results clearly indicate the superior predictive power of deep learning models, particularly those that combine spatial and temporal feature extraction. These models not only achieved lower error rates and higher R^2 values but also demonstrated resilience across seasonal variations and market disruptions. Furthermore, incorporating feature importance analysis and explainability tools such as SHAP values enhances transparency and user trust, addressing a key barrier to the adoption of AI tools in agriculture. The deployment of the models into a user-friendly digital dashboard further illustrates their real-world applicability, providing actionable insights to farmers, traders, and policymakers to improve market decisions, minimize post-harvest losses, and stabilize incomes. However, the study also acknowledges ongoing challenges, including data quality issues, infrastructure constraints in rural areas, the need for localized model tuning, and the dynamic nature of agricultural markets requiring periodic model updates. These limitations point toward future work that should focus on the incorporation of real-time IoT sensor data, satellite imagery, and integration with mobile-based platforms to enhance accessibility and responsiveness. Additionally, expanding the scope to include yield prediction, quality assessment, and risk modeling could support the development of a comprehensive AI-powered agricultural decision support system. Overall, this research affirms that AI-ML technologies, when thoughtfully designed and contextually deployed, hold transformative potential for enhancing transparency, efficiency, and sustainability in agri-horticultural markets, particularly in regions where farmers are highly vulnerable to market uncertainties. As the agricultural sector moves toward data-driven decision-making, the proposed approach lays the groundwork for scalable, intelligent systems that not only forecast prices but also empower stakeholders across the value chain with timely, reliable, and interpretable information.

REFERENCES

1. Reddy, C. N. K., & Murthy, G. V. (2012). Evaluation of Behavioral Security in Cloud Computing. *International Journal of Computer Science and Information Technologies*, *3*(2), 3328-3333.
2. Murthy, G. V., Kumar, C. P., & Kumar, V. V. (2017, December). Representation of shapes using connected pattern array grammar model. In *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)* (pp. 819-822). IEEE.
3. Krishna, K. V., Rao, M. V., & Murthy, G. V. (2017). Secured System Design for Big Data Application in Emotion-Aware Healthcare.
4. Rani, G. A., Krishna, V. R., & Murthy, G. V. (2017). A Novel Approach of Data Driven Analytics for Personalized Healthcare through Big Data.
5. Rao, M. V., Raju, K. S., Murthy, G. V., & Rani, B. K. (2020). Configure and Management of Internet of Things. *Data Engineering and Communication Technology*, 163.
6. Ramakrishna, C., Kumar, G. K., Reddy, A. M., & Ravi, P. (2018). A Survey on various IoT Attacks and its Countermeasures. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, *5*(4), 143-150.
7. Chithanuru, V., & Ramaiah, M. (2023). An anomaly detection on blockchain infrastructure using artificial intelligence techniques: Challenges and future directions–A review. *Concurrency and Computation: Practice and Experience*, *35*(22), e7724.
8. Prashanth, J. S., & Nandury, S. V. (2015, June). Cluster-based rendezvous points selection for reducing tour length of mobile element in WSN. In *2015 IEEE International Advance Computing Conference (IACC)* (pp. 1230-1235). IEEE.
9. Kumar, K. A., Pabboju, S., & Desai, N. M. S. (2014). Advance text steganography algorithms: an overview. *International Journal of Research and Applications*, *1*(1), 31-35.
10. Hnamte, V., & Balram, G. (2022). Implementation of Naive Bayes Classifier for Reducing DDoS Attacks in IoT Networks. *Journal of Algebraic Statistics*, *13*(2), 2749-2757.
11. Balram, G., Anitha, S., & Deshmukh, A. (2020, December). Utilization of renewable energy sources in generation and distribution optimization. In *IOP Conference Series: Materials Science and Engineering* (Vol. 981, No. 4, p. 042054). IOP Publishing.
12. Subrahmanyam, V., Sagar, M., Balram, G., Ramana, J. V., Tejaswi, S., & Mohammad, H. P. (2024, May). An Efficient Reliable Data Communication For Unmanned Air Vehicles (UAV) Enabled Industry Internet of Things (IIoT). In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-4). IEEE.
13. Mahammad, F. S., Viswanatham, V. M., Tahseen, A., Devi, M. S., & Kumar, M. A. (2024, July). Key distribution scheme for preventing key reinstallation attack in wireless networks. In *AIP Conference Proceedings* (Vol. 3028, No. 1). AIP Publishing.

14. Lavanya, P. (2024). In-Cab Smart Guidance and support system for Dragline operator.
15. Kovoor, M., Durairaj, M., Karyakarte, M. S., Hussain, M. Z., Ashraf, M., & Maguluri, L. P. (2024). Sensor-enhanced wearables and automated analytics for injury prevention in sports. *Measurement: Sensors*, *32*, 101054.
16. Rao, N. R., Kovoor, M., Kishor Kumar, G. N., & Parameswari, D. V. L. (2023). Security and privacy in smart farming: challenges and opportunities. *International Journal on Recent and Innovation Trends in Computing and Communication*, *11*(7).
17. Madhuri, K. (2023). Security Threats and Detection Mechanisms in Machine Learning. *Handbook of Artificial Intelligence*, *255*.
18. Reddy, B. A., & Reddy, P. R. S. (2012). Effective data distribution techniques for multi-cloud storage in cloud computing. *CSE, Anurag Group of Institutions, Hyderabad, AP, India*.
19. Srilatha, P., Murthy, G. V., & Reddy, P. R. S. (2020). Integration of Assessment and Learning Platform in a Traditional Class Room Based Programming Course. *Journal of Engineering Education Transformations*, *33*, 179-184.
20. Reddy, P. R. S., & Ravindranadh, K. (2019). An exploration on privacy concerned secured data sharing techniques in cloud. *International Journal of Innovative Technology and Exploring Engineering*, *9*(1), 1190-1198.
21. Raj, R. S., & Raju, G. P. (2014, December). An approach for optimization of resource management in Hadoop. In *International Conference on Computing and Communication Technologies* (pp. 1-5). IEEE.
22. Ramana, A. V., Bhoga, U., Dhulipalla, R. K., Kiran, A., Chary, B. D., & Reddy, P. C. S. (2023, June). Abnormal Behavior Prediction in Elderly Persons Using Deep Learning. In *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)* (pp. 1-5). IEEE.
23. Yakoob, S., Krishna Reddy, V., & Dastagiraiah, C. (2017). Multi User Authentication in Reliable Data Storage in Cloud. In *Computer Communication, Networking and Internet Security: Proceedings of IC3T 2016* (pp. 531-539). Springer Singapore.
24. Sukhavasi, V., Kulkarni, S., Raghavendran, V., Dastagiraiah, C., Apat, S. K., & Reddy, P. C. S. (2024). Malignancy Detection in Lung and Colon Histopathology Images by Transfer Learning with Class Selective Image Processing.
25. Dastagiraiah, C., Krishna Reddy, V., & Pandurangarao, K. V. (2018). Dynamic load balancing environment in cloud computing based on VM ware off-loading. In *Data Engineering and Intelligent Computing: Proceedings of IC3T 2016* (pp. 483-492). Springer Singapore.
26. Swapna, N. (2017). „Analysis of Machine Learning Algorithms to Protect from Phishing in Web Data Mining‟. *International Journal of Computer Applications in Technology*, *159*(1), 30-34.
27. Moparthi, N. R., Bhattacharyya, D., Balakrishna, G., & Prashanth, J. S. (2021). Paddy leaf disease detection using CNN.
28. Balakrishna, G., & Babu, C. S. (2013). Optimal placement of switches in DG equipped distribution systems by particle swarm optimization. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, *2*(12), 6234-6240.
29. Moparthi, N. R., Sagar, P. V., & Balakrishna, G. (2020, July). Usage for inside design by AR and VR technology. In *2020 7th International Conference on Smart Structures and Systems (ICSSS)* (pp. 1-4). IEEE.
30. Amarnadh, V., & Moparthi, N. R. (2023). Comprehensive review of different artificial intelligence-based methods for credit risk assessment in data science. *Intelligent Decision Technologies*, *17*(4), 1265-1282.
31. Amarnadh, V., & Moparthi, N. (2023). Data Science in Banking Sector: Comprehensive Review of Advanced Learning Methods for Credit Risk Assessment. *International Journal of Computing and Digital Systems*, *14*(1), 1-xx.
32. Amarnadh, V., & Rao, M. N. (2025). A Consensus Blockchain-Based Credit Risk Evaluation and Credit Data Storage Using Novel Deep Learning Approach. *Computational Economics*, 1-34.
33. Shailaja, K., & Anuradha, B. (2017). Improved face recognition using a modified PSO based self-weighted linear collaborative discriminant regression classification. *J. Eng. Appl. Sci*, *12*, 7234-7241.
34. Sekhar, P. R., & Goud, S. (2024). Collaborative Learning Techniques in Python Programming: A Case Study with CSE Students at Anurag University. *Journal of Engineering Education Transformations*, *38*.
35. Sekhar, P. R., & Sujatha, B. (2023). Feature extraction and independent subset generation using genetic algorithm for improved classification. *Int. J. Intell. Syst. Appl. Eng*, *11*, 503-512.
36. Pesaramelli, R. S., & Sujatha, B. (2024, March). Principle correlated feature extraction using differential evolution for improved classification. In *AIP Conference Proceedings* (Vol. 2919, No. 1). AIP Publishing.
37. Tejaswi, S., Sivaprashanth, J., Bala Krishna, G., Sridevi, M., & Rawat, S. S. (2023, December). Smart

Dustbin Using IoT. In *International Conference on Advances in Computational Intelligence and Informatics* (pp. 257-265). Singapore: Springer Nature Singapore.
38. Moreb, M., Mohammed, T. A., & Bayat, O. (2020). A novel software engineering approach toward using machine learning for improving the efficiency of health systems. *IEEE Access*, *8*, 23169-23178.
39. Ravi, P., Haritha, D., & Niranjan, P. (2018). A Survey: Computing Iceberg Queries. *International Journal of Engineering & Technology*, *7*(2.7), 791-793.
40. Madar, B., Kumar, G. K., & Ramakrishna, C. (2017). Captcha breaking using segmentation and morphological operations. *International Journal of Computer Applications*, *166*(4), 34-38.
41. Rani, M. S., & Geetavani, B. (2017, May). Design and analysis for improving reliability and accuracy of big-data based peripheral control through IoT. In *2017 International Conference on Trends in Electronics and Informatics (ICEI)* (pp. 749-753). IEEE.
42. Reddy, T., Prasad, T. S. D., Swetha, S., Nirmala, G., & Ram, P. (2018). A study on antiplatelets and anticoagulants utilisation in a tertiary care hospital. *International Journal of Pharmaceutical and Clinical Research*, *10*, 155-161.
43. Prasad, P. S., & Rao, S. K. M. (2017). HIASA: Hybrid improved artificial bee colony and simulated annealing based attack detection algorithm in mobile ad-hoc networks (MANETs). *Bonfring International Journal of Industrial Engineering and Management Science*, *7*(2), 01-12.
44. AC, R., Chowdary Kakarla, P., Simha PJ, V., & Mohan, N. (2022). Implementation of Tiny Machine Learning Models on Arduino 33–BLE for Gesture and Speech Recognition.
45. Subrahmanyam, V., Sagar, M., Balram, G., Ramana, J. V., Tejaswi, S., & Mohammad, H. P. (2024, May). An Efficient Reliable Data Communication For Unmanned Air Vehicles (UAV) Enabled Industry Internet of Things (IIoT). In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)* (pp. 1-4). IEEE.
46. Nagaraj, P., Prasad, A. K., Narsimha, V. B., & Sujatha, B. (2022). Swine flu detection and location using machine learning techniques and GIS. *International Journal of Advanced Computer Science and Applications*, *13*(9).
47. Priyanka, J. H., & Parveen, N. (2024). DeepSkillNER: an automatic screening and ranking of resumes using hybrid deep learning and enhanced spectral clustering approach. *Multimedia Tools and Applications*, *83*(16), 47503-47530.
48. Sathish, S., Thangavel, K., & Boopathi, S. (2010). Performance analysis of DSR, AODV, FSR and ZRP routing protocols in MANET. *MES Journal of Technology and Management*, 57-61.
49. Siva Prasad, B. V. V., Mandapati, S., Kumar Ramasamy, L., Boddu, R., Reddy, P., & Suresh Kumar, B. (2023). Ensemble-based cryptography for soldiers' health monitoring using mobile ad hoc networks. *Automatika: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, *64*(3), 658-671.
50. Elechi, P., & Onu, K. E. (2022). Unmanned Aerial Vehicle Cellular Communication Operating in Non-terrestrial Networks. In *Unmanned Aerial Vehicle Cellular Communications* (pp. 225-251). Cham: Springer International Publishing.
51. Prasad, B. V. V. S., Mandapati, S., Haritha, B., & Begum, M. J. (2020, August). Enhanced Security for the authentication of Digital Signature from the key generated by the CSTRNG method. In *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1088-1093). IEEE.
52. Mukiri, R. R., Kumar, B. S., & Prasad, B. V. V. (2019, February). Effective Data Collaborative Strain Using RecTree Algorithm. In *Proceedings of International Conference on Sustainable Computing in Science, Technology and Management (SUSCOM), Amity University Rajasthan, Jaipur-India*.
53. Balaraju, J., Raj, M. G., & Murthy, C. S. (2019). Fuzzy-FMEA risk evaluation approach for LHD machine–A case study. *Journal of Sustainable Mining*, *18*(4), 257-268.
54. Thirumoorthi, P., Deepika, S., & Yadaiah, N. (2014, March). Solar energy based dynamic sag compensator. In *2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCEE)* (pp. 1-6). IEEE.
55. Vinayasree, P., & Reddy, A. M. (2025). A Reliable and Secure Permissioned Blockchain-Assisted Data Transfer Mechanism in Healthcare-Based Cyber-Physical Systems. *Concurrency and Computation: Practice and Experience*, *37*(3), e8378.
56. Acharjee, P. B., Kumar, M., Krishna, G., Raminenei, K., Ibrahim, R. K., & Alazzam, M. B. (2023, May). Securing International Law Against Cyber Attacks through Blockchain Integration. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 2676-2681). IEEE.
57. Ramineni, K., Reddy, L. K. K., Ramana, T. V., & Rajesh, V. (2023, July). Classification of Skin Cancer Using Integrated Methodology. In *International Conference on Data Science and Applications* (pp. 105-118). Singapore: Springer Nature Singapore.

58. LAASSIRI, J., EL HAJJI, S. A. Ï. D., BOUHDADI, M., AOUDE, M. A., JAGADISH, H. P., LOHIT, M. K., ... & KHOLLADI, M. (2010). Specifying Behavioral Concepts by engineering language of RM-ODP. *Journal of Theoretical and Applied Information Technology*, *15*(1).
59. Prasad, D. V. R., & Mohanji, Y. K. V. (2021). FACE RECOGNITION-BASED LECTURE ATTENDANCE SYSTEM: A SURVEY PAPER. *Elementary Education Online*, *20*(4), 1245-1245.
60. Dasu, V. R. P., & Gujjari, B. (2015). Technology-Enhanced Learning Through ICT Tools Using Aakash Tablet. In *Proceedings of the International Conference on Transformations in Engineering Education: ICTIEE 2014* (pp. 203-216). Springer India.
61. Reddy, A. M., Reddy, K. S., Jayaram, M., Venkata Maha Lakshmi, N., Aluvalu, R., Mahesh, T. R., ... & Stalin Alex, D. (2022). An efficient multilevel thresholding scheme for heart image segmentation using a hybrid generalized adversarial network. *Journal of Sensors*, *2022*(1), 4093658.
62. Srinivasa Reddy, K., Suneela, B., Inthiyaz, S., Hasane Ahammad, S., Kumar, G. N. S., & Mallikarjuna Reddy, A. (2019). Texture filtration module under stabilization via random forest optimization methodology. *International Journal of Advanced Trends in Computer Science and Engineering*, *8*(3), 458-469.
63. Ramakrishna, C., Kumar, G. K., Reddy, A. M., & Ravi, P. (2018). A Survey on various IoT Attacks and its Countermeasures. *International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)*, *5*(4), 143-150.
64. Sirisha, G., & Reddy, A. M. (2018, September). Smart healthcare analysis and therapy for voice disorder using cloud and edge computing. In *2018 4th international conference on applied and theoretical computing and communication technology (iCATccT)* (pp. 103-106). IEEE.
65. Reddy, A. M., Yarlagadda, S., & Akkinen, H. (2021). An extensive analytical approach on human resources using random forest algorithm. *arXiv preprint arXiv:2105.07855*.
66. Kumar, G. N., Bhavanam, S. N., & Midasala, V. (2014). Image Hiding in a Video-based on DWT & LSB Algorithm. In *ICPVS Conference*.
67. Naveen Kumar, G. S., & Reddy, V. S. K. (2022). High performance algorithm for content-based video retrieval using multiple features. In *Intelligent Systems and Sustainable Computing: Proceedings of ICISSC 2021* (pp. 637-646). Singapore: Springer Nature Singapore.
68. Reddy, P. S., Kumar, G. N., Ritish, B., SaiSwetha, C., & Abhilash, K. B. (2013). Intelligent parking space detection system based on image segmentation. *Int J Sci Res Dev*, *1*(6), 1310-1312.
69. Naveen Kumar, G. S., Reddy, V. S. K., & Kumar, S. S. (2018). High-performance video retrieval based on spatio-temporal features. *Microelectronics, Electromagnetics and Telecommunications*, 433-441.
70. Kumar, G. N., & Reddy, M. A. BWT & LSB algorithm based hiding an image into a video. *IJESAT*, 170-174.
71. Lopez, S., Sarada, V., Praveen, R. V. S., Pandey, A., Khuntia, M., & Haralayya, D. B. (2024). Artificial intelligence challenges and role for sustainable education in india: Problems and prospects. *Sandeep Lopez, Vani Sarada, RVS Praveen, Anita Pandey, Monalisa Khuntia, Bhadrappa Haralayya (2024) Artificial Intelligence Challenges and Role for Sustainable Education in India: Problems and Prospects. Library Progress International*, *44*(3), 18261-18271.
72. Yamuna, V., Praveen, R. V. S., Sathya, R., Dhivva, M., Lidiya, R., & Sowmiya, P. (2024, October). Integrating AI for Improved Brain Tumor Detection and Classification. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1603-1609). IEEE.
73. Kumar, N., Kurkute, S. L., Kalpana, V., Karuppannan, A., Praveen, R. V. S., & Mishra, S. (2024, August). Modelling and Evaluation of Li-ion Battery Performance Based on the Electric Vehicle Tiled Tests using Kalman Filter-GBDT Approach. In *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)* (pp. 1-6). IEEE.
74. Sharma, S., Vij, S., Praveen, R. V. S., Srinivasan, S., Yadav, D. K., & VS, R. K. (2024, October). Stress Prediction in Higher Education Students Using Psychometric Assessments and AOA-CNN-XGBoost Models. In *2024 4th International Conference on Sustainable Expert Systems (ICSES)* (pp. 1631-1636). IEEE.
75. Anuprathibha, T., Praveen, R. V. S., Sukumar, P., Suganthi, G., & Ravichandran, T. (2024, October). Enhancing Fake Review Detection: A Hierarchical Graph Attention Network Approach Using Text and Ratings. In *2024 Global Conference on Communications and Information Technologies (GCCIT)* (pp. 1-5). IEEE.
76. Shinkar, A. R., Joshi, D., Praveen, R. V. S., Rajesh, Y., & Singh, D. (2024, December). Intelligent solar energy harvesting and management in IoT nodes using deep self-organizing maps. In *2024 International Conference on Emerging Research in Computational Science (ICERCS)* (pp. 1-6). IEEE.
77. Praveen, R. V. S., Hemavathi, U., Sathya, R., Siddiq, A. A., Sanjay, M. G., & Gowdish, S. (2024, October). AI Powered Plant Identification and Plant Disease Classification System. In *2024 4th*

International Conference on Sustainable Expert Systems (ICSES) (pp. 1610-1616). IEEE.

78. Dhivya, R., Sagili, S. R., Praveen, R. V. S., VamsiLala, P. N. V., Sangeetha, A., & Suchithra, B. (2024, December). Predictive Modelling of Osteoporosis using Machine Learning Algorithms. In *2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS)* (pp. 997-1002). IEEE.
79. Kemmannu, P. K., Praveen, R. V. S., Saravanan, B., Amshavalli, M., & Banupriya, V. (2024, December). Enhancing Sustainable Agriculture Through Smart Architecture: An Adaptive Neuro-Fuzzy Inference System with XGBoost Model. In *2024 International Conference on Sustainable Communication Networks and Application (ICSCNA)* (pp. 724-730). IEEE.
80. Praveen, R. V. S. (2024). *Data Engineering for Modern Applications*. Addition Publishing House.