Advanced International Journal of Multidisciplinary Research

E-ISSN: 2584-0487
editor@aijmr.com

Volume 4, Issue 2, March-April 2026

CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

SMART RURAL UTILITY CONSUMPTION ANALYSIS AND FORECASTING

Dr. M. Thejovathi[1], Mulpuru V N V CH Sree Samhitha[2], B. Sahithi[3], B. Sruthi[4]

[1]Associate Professor, CSE(AI&ML), Vignan's Institute of Management and Technology for Women, HYD, India,
[2,3,4]BTech 4th year Student, CSE(AI&ML), Vignan's Institute of Management and Technology for Women, Hyderabad, India.

Abstract:
Electricity and water are two resources for households, directly impacting quality SMS notifications to understand their consumption. While these provide basic information, they only reflect usage and do not guide families in planning for future demand or managing resources efficiently. This drawback often leads to sudden expenses, wastage, and difficulty in adopting sustainable practices. The key problem lies in the absence of affordable and accessible monitoring tools for rural households. While urban areas may benefit from IoT-enabled smart meters and advanced dashboards, such technologies are expensive and difficult to implement in rural contexts. As a result, rural families lack the means to analyse trends or forecast future consumption, creating a clear gap that needs a cost-effective and practical solution. This proposal aims to design and develop an AI-based web application for Indian rural households. The system will be designed to process existing bills and SMS updates, extracting essential details such as units consumed, billing cycles, and charges through Optical Character Recognition (OCR). Historical records up to seven years will be analysed using hybrid machine learning and ensemble learning approaches to generate accurate forecasts of future consumption and expected costs. The interface will be kept simple and accessible, with potential for local language support, ensuring usability even for rural families. The expected outcome is a low-cost and user friendly tool that will help rural households visualize their usage patterns, forecast future requirements weeks or months in advance, and adopt better resource management practices. By turning static bills into predictive insights, the proposed solution has the potential to improve planning, reduce wastage, promote sustainable living, and contribute to national goals of energy efficiency.

Keywords: Indian Rural Households, Utility Consumption, Forecasting, AI, Hybrid Machine Learning, Ensemble Learning, Sustainability.

I. INTRODUCTION:

Efficient management of electricity and water resources is essential for sustainable development and financial stability, particularly in rural communities. While urban regions increasingly adopt IoT-enabled smart meters and real-time analytics platforms, rural households largely rely on conventional monthly billing systems that provide only historical consumption data. These traditional methods lack predictive insights, making it difficult for users to anticipate future usage patterns or manage expenses effectively. Moreover, IoT-based solutions often involve high installation costs and infrastructure requirements, limiting their feasibility in rural environments. To address this gap, this paper proposes a Smart Rural

E-ISSN: 2584-0487
editor@aijmr.com
Volume 4, Issue 2, March-April 2026
CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

Utility Consumption Analysis and Forecasting framework that transforms static billing data into predictive intelligence using Artificial Intelligence. Weather-integrated AI models have been applied for energy demand prediction and cost estimation
[1], while user-centric feedback interfaces have been shown to improve residential energy awareness and monitoring [2]. The system utilizes Optical Character Recognition (OCR) to extract consumption details from uploaded electricity and water bills and integrates weather parameters to enhance forecasting accuracy. Hybrid time-series models combining ARIMA with deep learning approaches have demonstrated strong performance in electricity demand forecasting [3], and ensemble and gradient boosting methods have achieved higher accuracy in multivariate water demand prediction when multiple environmental factors are considered [4]. Studies on household water consumption further highlight the influence of behavioural and environmental variables on resource usage patterns [5]. Recent advancements in hybrid machine learning, deep learning, and big data analytics further validate the effectiveness of AI-based forecasting frameworks for complex consumption prediction tasks [6][7][8][9][10]. By eliminating hardware dependency and incorporating multilingual accessibility, the proposed solution provides a cost-effective, scalable, and user-centric approach to intelligent rural utility management.

II. RELATED WORK:

Recent research shows that machine learning techniques significantly improve forecasting accuracy in utility consumption systems. Weather-integrated artificial intelligence models have been applied for energy demand prediction and cost estimation [1], while user-centric energy feedback interfaces improve residential energy awareness and monitoring [2]. Hybrid time-series models combining ARIMA with deep learning approaches such as LSTM have demonstrated strong performance in electricity demand forecasting [3]. Ensemble and gradient boosting methods have also been applied for multivariate water demand prediction, achieving higher accuracy when multiple environmental factors are considered [4]. Studies on household water consumption further highlight the influence of behavioural and environmental variables on resource usage patterns [5]. Recent advancements explore hybrid machine learning and time-series forecasting techniques such as XGBoost and SARIMA for improving demand prediction performance [6], whereas deep learning models are increasingly used for weather forecasting and temporal pattern analysis [7]. Comparative studies on gradient boosting models demonstrate their effectiveness in handling nonlinear consumption data [8]. Hybrid deep learning architectures and big data analytics frameworks have also been proposed for predictive decision-support systems [9], [10]. However, many existing solutions depend on urban datasets or complex infrastructure, limiting their applicability in rural environments. The proposed system addresses these limitations by providing a cost-effective AI-based forecasting framework tailored for rural utility consumption analysis.

III. PROPOSED SYSTEM:

A. Overview of Proposed System:

The proposed Smart Rural Utility Consumption Analysis and Forecasting system is a web-based AI framework designed to assist rural households in analysing and predicting electricity and water usage. The system eliminates the need for IoT hardware by utilising uploaded utility bills as the primary data source. It integrates OCR-based data extraction, weather-enhanced feature processing, and hybrid machine learning models to generate accurate consumption forecasts and cost estimations. The system ensures multilingual accessibility and user-friendly visualisation to support rural communities.

E-ISSN: 2584-0487
editor@aijmr.com
Volume 4, Issue 2, March-April 2026
CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

B. System Architecture:
The architecture consists of five core modules:
• User Authentication & Language Module
• Bill Processing & OCR Module
• Data Management & Storage Module
• Hybrid ML Forecasting Module
• Visualization & Recommendation Module

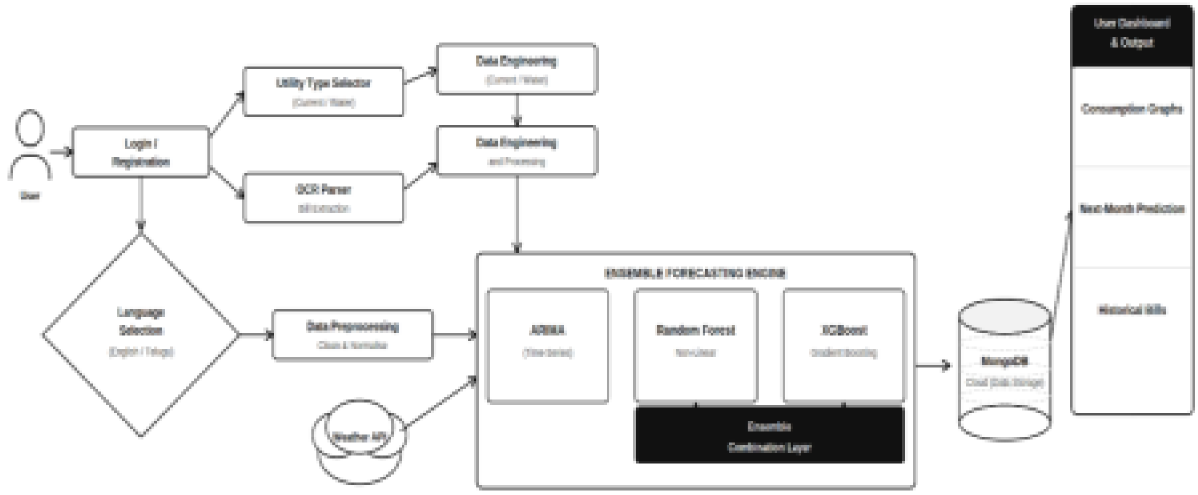


Figure 1: System Architecture for Smart Rural Utility Composition Analysis & Forecasting

• **User Authentication & Language Module:** This module manages user registration, login, and language selection (Telugu/English). The selected language preference is stored in the user profile. The entire system interface dynamically adapts based on this preference. This ensures accessibility and ease of use.
• **Bill Processing & OCR Module:** Users upload electricity and water bills separately. The system performs OCR to extract consumption units, billing period, and total amount. Extracted data is validated and structured. This replaces manual data entry and reduces errors.
• **Data Management Module:** Structured consumption records are stored in a database. Historical data is maintained for both electricity and water independently. This module supports retrieval for visualization and forecasting. Secure storage ensures data consistency.
• **Hybrid ML Forecasting Module:** The forecasting engine combines ARIMA, Random Forest, and XGBoost models. Time-series patterns and nonlinear behaviour are captured effectively. Weighted ensemble prediction improves forecasting accuracy. The model outputs predicted units and estimated cost.
• **Visualization & Recommendation Module:** This module displays consumption graphs and forecast results separately for electricity and water. It provides estimated billing amounts and trend analysis. Smart recommendations are generated based on prediction patterns. Optional voice output enhances usability
.

IV. IMPLEMENTATION:

A. Development Environment

The system is implemented as a full-stack web application. The frontend is developed using React.js to provide an interactive and multilingual user interface. The backend is built using Python with Flask, which

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 4, Issue 2, March-April 2026
CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

handles authentication, OCR processing, data management, and machine learning execution. Database (MongoDB) is used to store user profiles, extracted bill data, and forecast results. The modular design supports scalability and easy maintenance.

B. User Authentication and Language Handling

During registration, users select their preferred language (Telugu or English). The selected language is stored in the user profile and applied dynamically across all pages using conditional rendering. Secure login authentication is implemented to ensure data privacy. Session management maintains user-specific dashboard access and historical records.

C. Bill Processing and OCR Implementation

Electricity and water bills are uploaded separately in image or PDF format. The system uses Tesseract OCR to extract relevant textual data such as units consumed, billing period, and total amount. Image preprocessing techniques including grayscale conversion and noise filtering improve extraction accuracy. The extracted text is parsed and converted into structured database records.

D. Data Preprocessing and Storage

Extracted consumption values are cleaned and formatted into time-series structure. Historical data is maintained separately for electricity and water usage. Necessary preprocessing such as normalization and date formatting is performed before model input. All processed records are securely stored for forecasting and visualization.

E. Hybrid Machine Learning Forecasting

The forecasting module integrates ARIMA, Random Forest Regressor, and XGBoost Regressor models. ARIMA captures trend and seasonal components, while Random Forest and XGBoost model nonlinear consumption behaviour. Predictions from all models are combined using weighted averaging to improve accuracy. The final output includes forecasted units and estimated billing cost.

F. Visualization and Recommendation System

Consumption trends and forecast results are displayed using graphical charts on the dashboard. Electricity and water data are visualized separately for clarity. Based on predicted patterns, the system generates usage recommendations in the selected language. Voice output functionality is implemented to enhance accessibility for rural users.

V. ALGORITHM:

INPUT: User credentials, Bill image, Historical consumption, Weather data
OUTPUT: Forecasted units, estimated cost, Smart recommendations

```
BEGIN
 INITIALIZE database, OCR engine, ML models, Weather API
 IF new user THEN
 STORE (username, HASH(password), language_preference) -> database  ELSE
 VERIFY credentials
 GENERATE JWT token
 LOAD dashboard in selected language
 END IF
 ACCEPT bill image (electricity or water)
 preprocessed_image <- GRAYSCALE + DENOISE + THRESHOLD(image)   raw_text <- TESSERACT_OCR(preprocessed_image)
 units_consumed <- EXTRACT units from raw_text
 billing_period <- EXTRACT date from raw_text
```

E-ISSN: 2584-0487
editor@aijmr.com
Volume 4, Issue 2, March-April 2026
CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

total_amount <- EXTRACT amount from raw_text
STORE extracted data -> database
historical_data <- FETCH all bills for user from database
REMOVE invalid records
SORT by billing date
FILL missing values using linear interpolation
normalized_data <- MIN_MAX_NORMALIZE(historical_data)
weather <- FETCH (temperature, humidity, rainfall) from API
feature_set <- MERGE (normalized_data, weather)
arima_pred <- ARIMA_MODEL.fit(historical_data).predict(next_month) rf_pred <- RF_MODEL.predict(feature_set)
xgb_pred <- XGB_MODEL.predict(feature_set)
forecast <- (0.4 x arima_pred) + (0.3 x rf_pred) + (0.3 x xgb_pred)
estimated_cost <- forecast x rate_per_unit
IF trend > 15% increase THEN ADD alert recommendation END IF
IF temperature > 35 degrees THEN ADD cooling tip END IF
IF forecast > threshold THEN ADD energy saving tips END IF
IF language == Telugu THEN TRANSLATE recommendations END IF
DISPLAY graphs, forecast, cost, recommendations
IF voice requested THEN
CLEAN text
GENERATE audio <- gTTS(clean_text, language)
PLAY audio
END IF
bills <- FETCH bills WHERE user_id = current_user ORDER BY date DESC DISPLAY consumption trend chart
DISPLAY cost analysis chart
END

VI. RESULTS:

The system successfully processed uploaded bills and extracted data with high accuracy using OCR. Forecast predictions demonstrated reliable short-term estimation performance. Visualization dashboards clearly presented consumption trends and predicted costs. Experimental evaluation confirmed that the hybrid model produced higher accuracy than individual models.

AIJMR

Advanced International Journal of Multidisciplinary Research

E-ISSN: 2584-0487

editor@aijmr.com

Volume 4, Issue 2, March-April 2026

CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277

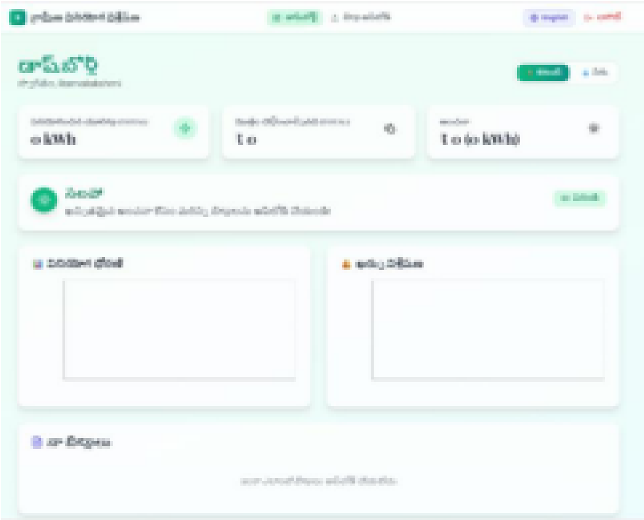


Figure 2: Display Dashboard **Figure 3:** Bill Upload

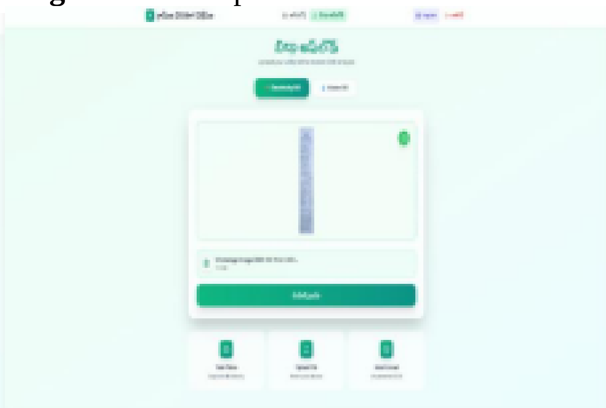
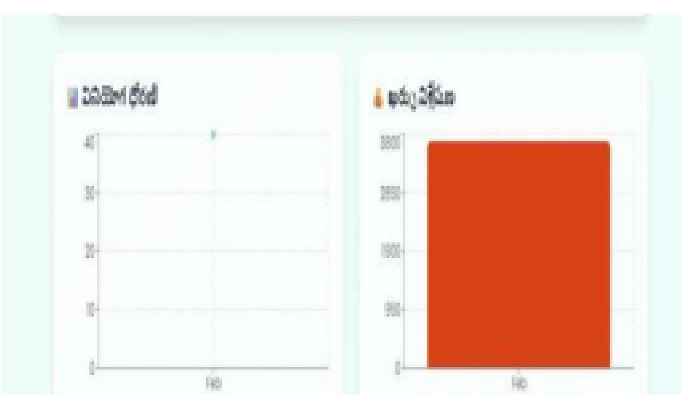


Figure 4: OCR Extraction of Electricity Bill **Figure 5:** OCR Extraction of Water Bill



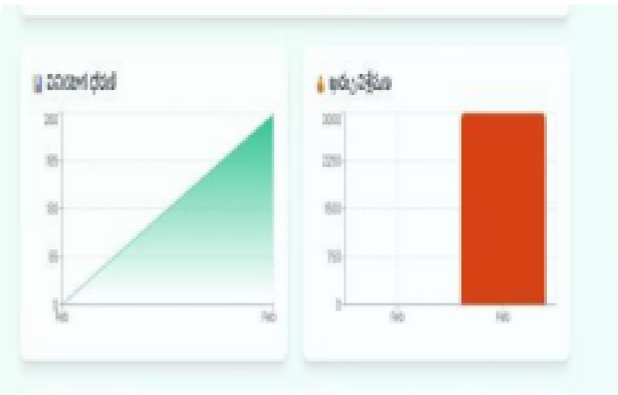


Figure 6: Electricity Consumption Graph **Figure 7:** Water Consumption Graph

Advanced International Journal of Multidisciplinary Research
E-ISSN: 2584-0487
editor@aijmr.com
Volume 4, Issue 2, March-April 2026
CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277


Figure 8: Display Usage, amount and forecast **Figure 9:** Display Usage, amount and forecast (Electricity) (Water)

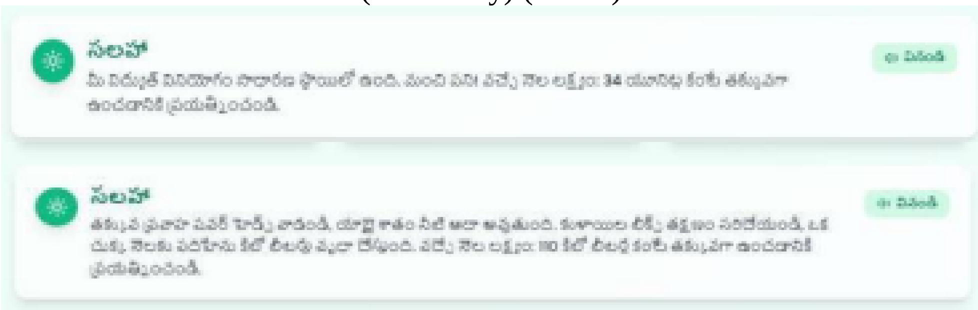


Figure 10,11: Display Recommendations

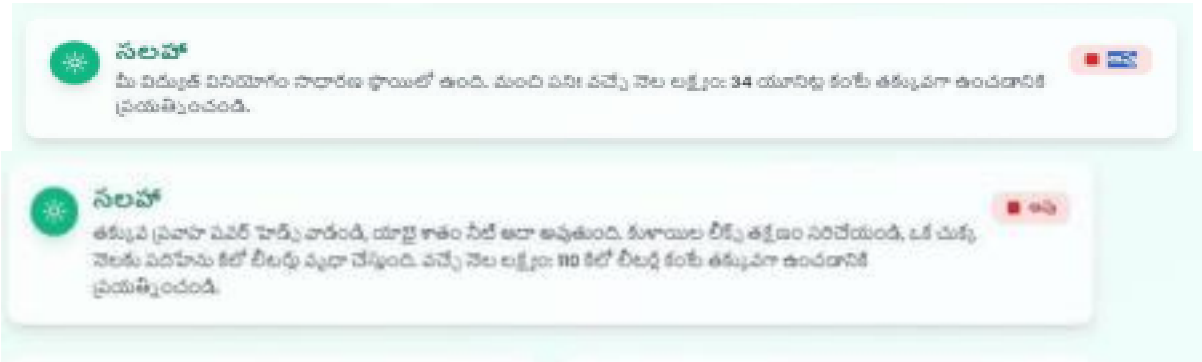


Figure 12,13: Play Voice Recommendation

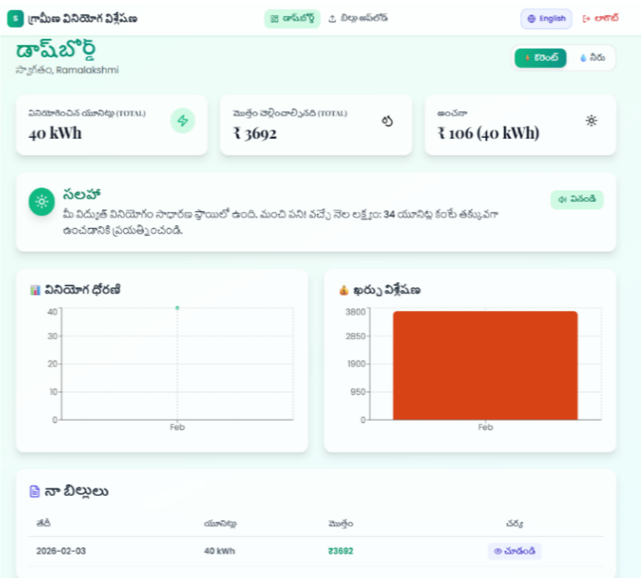

Figure 14: View previous bills and consumption **Figure 15:** View Previous bills and consumption (Electricity) (Water)

Advanced International Journal of Multidisciplinary Research

E-ISSN: 2584-0487
editor@aijmr.com

Volume 4, Issue 2, March-April 2026

CrossRef DOI: 10.62127/aijmr.2026.v04i02.1277



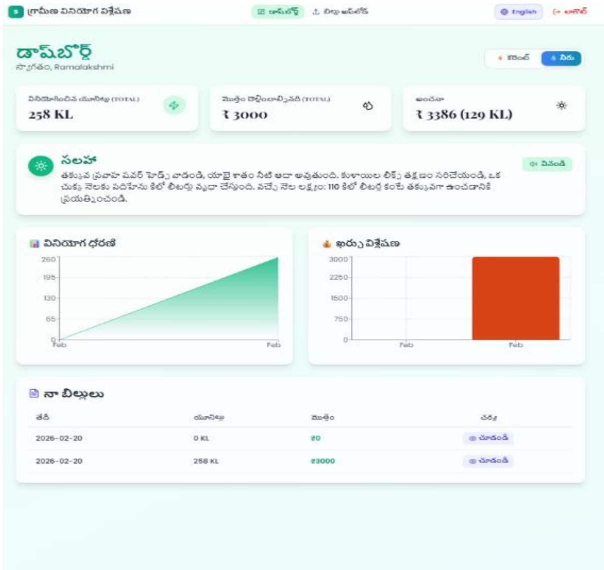


Figure 16: Dashboard display (Electricity) **Figure 17:** Dashboard display (Water)

VII. CONCLUSION:

This research presents an intelligent, low-cost forecasting platform that empowers rural households to manage electricity and water consumption efficiently. By integrating OCR, hybrid machine learning, and weather data, the system converts traditional billing information into predictive insights. The framework promotes sustainability, financial planning, and resource efficiency while remaining accessible to rural communities without requiring specialised hardware**.**

REFERENCES:

[1] V. Raji and E. Rajendran, "Weather-Integrated Smart Energy Administration Forecasting and Cost Estimation Powered in Artificial Intelligence," WSEAS Transactions on Environment and Development, 2025.

[2] Madhur Garg, Vishal Garg, Priyanka Srivastava, and Rishika Agarwal, "Interface Design for Residential Energy Feedback in the Indian Context," Energy Informatics, SpringerOpen, 2023. [3] C. Brindha, S. Sivakumar, K. Rajakali, and G. Pandiya Rajan, "Smart Electricity Demand Forecasting Using Hybrid ARIMA Model with LSTM Algorithm," International Journal of Creative Research Thoughts (IJCRT), 2022.

[4] Paul Banda, Muhammed A. Bhuiyan, Kevin Zhang, and Andy Song, "Multivariate Monthly Water Demand Prediction Using Ensemble and Gradient Boosting Machine Learning Techniques," Proceedings of the International Conference on Evolving Cities (ICEC), 2023.

[5] Cominola et al., "The Determinants of Household Water Consumption: A Review and Assessment Framework for Research and Practice," Environmental Modelling & Software, Elsevier. [6] Thejovathi, M., ChandraSekharaRao, M.V.P., Priyadharsini, E.J., Siddi, S., Karthik, B. and Abbas, S.H., 2024, November. Optimizing product demand forecasting with hybrid machine learning and time series models: A Comparative Analysis of XGBoost and Sarima. In Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024). [7] Johri, S., Divyajyothi, M.G., Anitha, S., Rani, M.S., Murari, T. and Shirisha, N., 2023,

November. A novel deep learning approach for capturing time series dependencies and improving short-term weather forecasting. In 2023 Seventh International Conference on Image Information Processing (ICIIP) (pp. 357-362). IEEE.
[8] Thejovathi, M. and Rao, D.M.C.S., 2024. Evaluating the performance of xgboost and gradient boost models with feature extraction in fmcg demand forecasting: A feature-enriched comparative study. J. Theor. Appl. Inf. Technol, 102, pp.4158-4163.
[9] Acharjee, P.B., Magadum, A.A., Thejovathi, M., Jain, R., Umarani, K. and Nishant, N., 2023, December. An innovative method for election prediction using hybrid a-bicnn-rnn approach. In 2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS) (pp. 765-770). IEEE.
[10] Murari, T., Prathiba, L., Singamaneni, K.K., Venu, D., Nassa, V.K., Kohar, R. and Uparkar, S.S., 2022. Big data analytics with OENN based clinical decision support system., *31*(2), pp.1241-1256.