AI Enhanced Supply Chain Management for Tea Leaves in Agriculture

2024/25 - 303

Bandara .D.M.A.T.D

B.Sc. (Hons) Degree in Information Technology Specialization in Information Technology

Department Of Information Technology

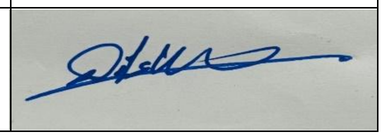
Sri Lanka Institute of Information Technology

Sri Lanka

August 2024

Declaration

I declare that this is my own work and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student ID	Signature
Bandara.D.M.A.T.D	IT21372162	

The above candidates are carrying out research for the undergraduate dissertation under my supervision.

……………………………..   ……………………

Signature of the Supervisor: Date:

(Mr.Uditha Dharmakeerthi)

ACKNOWLEDGMENT

I would like to express my deepest gratitude to my research supervisor Mr.Uditha Dharmakeerthi and my co-supervisor Ms. Lokesha Weerasinghe for their guidance, patience, and support throughout my research journey. Their expertise and insightful feedback have been invaluable in shaping our research. I would also like to thank my teammates for their collaborative efforts to make this project a success. Finally, my heartfelt gratitude goes out to my family and friends for always encouraging me and keeping me strong in times that were challenging throughout the course of this project.

ABSTRACT

For industries like tea production, which are impacted by labor, environmental, and logistical factors, the effectiveness of agricultural supply chains is essential to guaranteeing sustainability and profitability. In order to optimize important supply chain elements at the Watawala Tea Factory in Sri Lanka, this study suggests combining artificial intelligence (AI) and predictive analytics. Weather-based demand forecasting, inventory control, supply chain risk reduction, and logistics optimization are the four primary areas of focus for this study. The project creates predictive models to support data-driven decision-making, reduce waste, and guarantee product availability by utilizing historical weather, labor, and traffic datasets in addition to domain-specific insights. While inventory models convert monthly demand forecasts into weekly material requirements, the demand forecasting system ties synthetic sales data with weather patterns. Labor data is used in risk models to project shortages; traffic patterns are predicted in logistics models to maximize delivery plans. These elements taken together seek to improve operational resilience, lower delays, and build a more open, flexible tea supply chain.

Keywords: supply chain optimization, predictive analytics, demand forecasting, inventory management, logistics, labor risk mitigation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBRIVATIONS

Abbreviation	Full Form
AI	Artificial Intelligence
LSTM	Long Short-Term Memory
ML	Machine Learning
SCM	Supply Chain Management
DL	Deep Learning
UI	User Interface
API	Application Programming Interface
SA	State A
SB	State B
SC	State C

1.INTRODUCTION

1.1 Background & Literature Survey

Tea business contributes significantly to the agrarian economies of most countries, like Sri Lanka, where it is among the primary sources of jobs and export revenue. Tea farming is a labor-intensive process, particularly in plucking and early processing, where manual handling is required to achieve quality. However, of late, the industry has been facing an increasingly deteriorating labor shortage. Rural-urban migration, aging workforce, lack of incentives, and uncertainty of seasonality have consolidated the demand-supply gap of labor. Operational inefficiency, missed harvesting opportunities, and massive yield loss and degradation of quality have been the consequences.

In light of this, the introduction of Artificial Intelligence (AI) to the farm supply chain represents a new way of mitigating interruptions caused by labor. This study is interested in applying AI in the making of labor forecasts for the tea supply chain and proposing a forecasting model with Long Short-Term Memory (LSTM) neural networks. LSTM models are able to learn long-term dependencies from sequential data and hence are best suited for applications of time-series forecasting such as forecasting labor availability based on past trends, weather data, and market-driven demand signals.

The final goal of this research is to develop an AI-strengthened decision support system for decision-making that accurately forecasts labor availability in different tea-growing regions. Through this, it aims to help tea estate managers and supply chain planners optimize the distribution of workforce, reduce wastage due to under-harvesting, and provide continuity to supply chain operations. The system will be adaptive and scalable, capable of incorporating region-specific variables and real-time environmental inputs. This intelligent forecasting mechanism will allow stakeholders to plan labor-intensive activities more effectively and enable smoother coordination between field operations and downstream processing facilities.

Moreover, the system to be adopted also has a forecasting model that not only predicts labor availability but also determines whether there will be a surplus or shortage of workforce against the predicted demand. This surplus-shortage determination is particularly significant in optimizing resource allocation and managing operational risks. The use of real-time data storage and retrieval through a cloud-integrated platform also enhances the functionality of the model for field managers and decision-makers. As the tea business evolves to incorporate the implementation of digital transformation, this research comes with the general objective of smart agriculture through solving a critical but less-researched area in supply chain management.

In the environment of modern agriculture, the intersection of digital technology advanced more rapidly with the growing trend of Precision Agriculture and Industry 4.0 programs. The majority of innovations have targeted yield estimation, monitoring of soil condition, automatic irrigation, and pest management. In spite of the significance of labor availability an essential success variable for conducting agricultural activities, little has been researched from data-driven and predictive analytics viewpoints. All the traditional methods rely on heuristic estimates or static historical data, and they fail to respond adequately to existing environmental fluctuations and uncertain socio-economic alterations.

In the recent past, machine learning has been adopted in agriculture with Support Vector Machines (SVM), Decision Trees, and Random Forests being used in classification and prediction. These models have, however, shown inadequacies in handling temporal dependencies in time-series data, which is an aspect important to the understanding of labor trends of availability. Of the deep learning techniques, Long ShortTerm Memory (LSTM) networks have been highly effective in sequence modeling applications due to their ability to learn long-range historical patterns and make multi-step predictions. Studies have proved that LSTM models outperform traditional statistical approaches like ARIMA in managing sophisticated, nonlinear, and multivariate farm forecasting tasks.

True as their potential, however, the use of LSTM and similar models has barely been noticed in agricultural labor forecasting, and few have been specially tuned to the specific rhythms of tea farming. The specificity of tea harvesting—where picking cycles are brief, demand is seasonally sensitive, and expertise in the labor force is vital—demands a more specific and more effective forecasting approach. In addition, current research tends to generalize agricultural issues without considering regional differences and crop-specific needs, leading to inapplicable or misleading models for high-value, labor-intensive crops such as tea.

Another key missing link in the existing literature is a failure to integrate correlation between labor estimates and external factors such as weather, rainfall, and market demand. Tea agriculture is highly susceptible to climate conditions that have a direct impact on the growth of tea leaves and harvest timing. This research integrates meteorology and labor statistics to improve predictability and enable anticipatory planning for the labor force. Besides, the work already available tends to ignore regional forecasting. Most models operate at national or global scales, whereas labor availability may vary drastically from one tea-growing region to another depending upon local climatic, economic, and demographic situations.

This therefore attempts to close these gaps through the development of a region-based, LSTM-activated labor prediction model that accepts both demand-oriented (e.g., market orders) and supply-oriented (e.g., weather and historical patterns of labor) inputs. To this end, it not only contributes to academic literature on AI in agriculture but also offers an actual solution to tea supply chain stakeholders to closely align labor supply with production schedules. This conforms to the general goals of sustainable agriculture practice, digital expansion, and supply chain resilience in emerging economies where tea production is a significant sector.

1.2 Research Gap

Despite growing interest in digitalization and smart agriculture in farm practices, the issue of labor scarcity in the tea industry is still untouched to a large extent, particularly from a predictive and data-driven approach. While past research in agriculture has generally been positive about the application of technological innovations for monitoring crop health, irrigation control, and yield prediction, relatively few studies have explored the field of labor availability forecasting especially with the accuracy and adaptability needed in labor-intensive crops like tea. Most existing approaches rely heavily on historical labor data or broad trend analysis, often overlooking the nuanced, dynamic variables affecting labor availability in the field. These methods are not reactive enough to cope with unexpected variables such as seasonal migration, changing climatic patterns, and arbitrary market fluctuations, all of which directly affect labor demand and supply within a region.

Furthermore, although usual machine learning models such as Decision Trees and Random Forest have previously been applied in agricultural prediction exercises, they do not perform so well in representing temporal dependencies an important element in learning labor patterns that change with time. Deep learning algorithms like Long Short Term Memory (LSTM) networks, specifically designed for time-series forecasting, have shown immense potential in other areas of agriculture and logistics but yet to be unleashed to their maximum potential in labor forecasting for tea cultivation. Even within the thin literature employing LSTM or similar models, no regionalization or crop-specific variables are accommodated, and existing solutions are too generalized to be efficiently deployed within the tea industry. So localized is the tea cultivation process where the climatic conditions, manpower demography, and picking cycles vary enormously between estates that the tea industry needs a more expert and scenario-based approach.

Moreover, current systems do not include external driving variables such as rainfall quantities, temperature fluctuations, and export market demand surges into labor forecasting models. All these play pivotal roles in planning work-intensive operations like plucking, which not only has a time constraint but also hinges on the availability of trained manpower. Absence of such multi dimensional forecasting renders labor management systems less feasible and reliable in practical use. There is also considerable absence of equipment that monitors labor excess or deficit in relation to projected raw material needs, making estate managers have no decision-making information to manage workforce capacity versus operational demand.

This research tries to address these significant gaps by developing an AI powered labor forecasting tool that leverages the temporal modeling potential of LSTM networks, incorporates weather and market data, and presents region-specific results for tea estate planning. With the intersection of labor dynamics and supply chain enhancement, this study strives toward a central but neglected dimension of agricultural efficiency, with implications for academic literature as well as the operational resilience of the tea industry.

Features	Research 1	Research 2	Research 3	Proposed System
AI-Driven SCRM in Tea Industry	✘	✓	✘	✓
Data-Driven Risk Assesment Models	✘	✘	✓	✓
Sustainability	✓	✘	✓	✓
Real-Time Risk Monitoring Systems	✘	✓	✘	✓
Web Application	✓	✓	✘	✓
Tailored AI Models for Tea Supply Chain Characteristics	✘	✘	✘	✓

Table 1.1.1 Comparison with Previous Research

Secondly, while AI-based solutions have been proposed to address labor problems in tea estates, such as creating AI-based harvesting machines that can observe in real time and pluck exactly, these technologies entail a significant amount of capital investment and may not be accessible to small farmers. Also, these solutions are focused more on mechanization rather than predictive workforce planning, and there is a gap left by not going in for proactive human resource planning and allocation.

Additionally, recent studies on supply chain resilience in the tea industry have centered on the requirement for collaboration and information sharing among stakeholders to mitigate the effects of disruptions. None of these studies specifically addresses the integration of predictive analytics for labor forecasting as being part of the supply chain management process.

Concisely, there is a clear research gap in the development and implementation of AI-driven predictive models, specifically with LSTM networks, in forecasting labor availability in the tea industry. Bridging the gap can contribute significantly to enhancing supply chain resilience and efficiency through the implementation of forward-thinking labor management strategies in accordance with the unique demands of tea cultivation.

1.3 Research Problem

Watawala, located in the central highlands of Sri Lanka, is one of Sri Lanka's most famous tea-growing regions, yielding some of the world's finest Ceylon tea. The prosperity of tea cultivation here, however, is highly dependent on the availability and timely utilization of labor, particularly for plucking and early stage processing operations. In recent years, though, tea estates began to face an increasingly critical problem: a chronic shortage of available labor. This labor shortage is due to several region specific dynamics, including rural-urban migration, aging worker populations of estates, and insufficient interest in plantation work among the younger generation. As a result, tea producers are struggling to balance labor supply with the fluctuating demands of production, especially during peak harvesting periods. The imbalance has led to under-harvesting, quality degradation, delayed shipments, and inefficiencies in local supply chain operations.

Despite the growing importance of digital agriculture, labor planning at Watawala is still largely manual and reactive, relying on past trends and estimates that do not factor in real-time field conditions or developing socio-economic trends. These antiquated methods are no longer sufficient in an environment where both labor dynamics and climatic conditions are changing at a rapid pace. Furthermore, there is a notable lack of smart systems capable of forecasting short-term labor availability by taking into consideration the important local variables such as rainfall patterns, estate-specific labor patterns, and market-driven production planning. This inability to forecast hinders the ability of estate managers and supply chain operators to anticipate and prepare for labor shortages or surpluses, leading to resource allocation inefficiencies and lost economic value.

Existing literature on AI applications in agriculture is more focused on areas like crop yield prediction, soil health monitoring, and pest detection with little attention to human labor forecasting a critical input in the tea industry. In addition, most research is generalized and does not offer the region specific data that is necessary in areas like, where labor trends and environmental conditions are very distinct from other tea-growing regions. The lack of such detailed, localized models also limits the effectiveness of any AI-driven planning tools in real-world estate operations.

This research addresses the critical problem of labor forecasting in Watawala by developing an AI-driven solution tailored to the region's unique agricultural landscape. Using Long Short Term Memory (LSTM) neural networks, the system learns from the past labor availability data, local weather patterns, and demand forecasts to predict the labor supply for short-term time periods. The model captures the real operational dynamics of estates and, therefore, can make accurate labor forecasts for different zones (State A, B, and C of Watawala). Among the most significant innovations of this solution is the real-time analysis of labor surplus or deficit versus expected demand, giving stakeholders a decision support system capable of guiding labor deployment, contract planning, and supply chain coordination.

Through the integration of this forecasting model into a cloud-based system, the solution delivers accessibility and usability for territorial authorities, planners, and estate managers. Lastly, the research solves the overall issue of labor uncertainty by rendering labor availability a measurable, predictable, and actionable aspect of supply chain management. The research fills a significant gap in theory and practice by presenting a locally tailored, AI powered solution to reduce labor shortages in one of Sri Lanka's most prominent tea-growing areas.

1.4 Research Objectives

The overall objective of this research is to develop an AI-powered prediction system capable of effectively forecasting short term labor availability in the Watawala region tea supply chain in Sri Lanka. By alleviating the principal issue of labor shortages, the research aims to improve decision-making and operational efficiency in tea estate management through the application of advanced machine learning techniques specifically, Long Short Term Memory (LSTM) neural networks. They are designed to capture temporal dependencies between historical labor patterns, local weather information, and production demand volatility to forecast labor availability over key estate areas for a week horizon. In doing so, the study seeks to bridge the gap between reactive labor planning and intelligent, data-driven forecasting.

Another primary objective is to analyze the relationship between forecasted labor supply and estimated demand for raw material (tea leaves) in the same region. Based on the AI-predicted labor availability and its comparison with the estimated labor requirements calculated by applying demand multipliers and the normal plucking capacities the model calculates the period of labor surplus or deficit. This insight is critical in enabling forward-planning resource allocation, preventing losses due to delayed harvesting, and maintaining supply chain continuity. The system also automates data analysis and stores findings in a centralized database, providing a unified digital solution that is accessible to planners and estate managers through a web based interface.

Furthermore, the study aims to create a localized, scalable forecasting framework that reflects the unique agricultural and socio-economic circumstances of Watawala. Unlike generic AI applications meant for more generic agricultural environments, this study tackles location-specific problems such as weather effect, estate specific work patterns, and variable seasonal cycles of production. By doing this, it not only contributes to the scholarly body of work on AI enhanced agricultural supply chains but also offers a functional tool with on-ground relevance. The ultimate goal is to reduce the negative impacts of labor shortages in tea cultivation while introducing a sustainable, intelligent solution for supply chain planning in Sri Lanka's plantation sector.

1.4.1 Main Objective

Creating a localized AI-driven labor forecasting model that can precisely predict the short-term labor availability in the tea supply chain in the Watawala region of Sri Lanka is the primary goal of this study. The model attempts to proactively address the issues brought on by fluctuating labor supply by utilizing Long Short-Term Memory (LSTM) neural networks and combining historical labor data with weather and production demand patterns. In one of Sri Lanka's most important tea-producing regions, the ultimate objective is to improve operational efficiency, decrease harvest delays, and assist tea estate managers in making data-driven decisions. This will guarantee sustainable and continuous supply chain performance.

1.4.2 Specific Objectives

- **Data Collection and Preprocessing**:
  - Gather historical labor data, weather conditions, and regional demand data specific to the Watawala region.
  - Preprocess and normalize the data to ensure consistency and usability for modeling.
- **Development of LSTM Model**:
  - Design and train a Long Short-Term Memory (LSTM) model to forecast labor availability based on historical labor trends and environmental factors.
  - Tailor the model to handle seasonal and regional variations specific to the tea industry in Watawala.
- **Integration of Real-Time Data**:
  - Incorporate real-time weather and market demand data to dynamically adjust labor predictions.
  - Consider precipitation, temperature, and tea production demand fluctuations in the forecasting model.
- **Decision Support Tool Development**:
  - Build a user-friendly tool to display labor availability forecasts and excess-deficiency analysis.
  - Enable stakeholders to make informed decisions on labor allocation and optimize workforce distribution.

- **System Evaluation**:
  - Evaluate the AI-based labor forecasting system through case studies or simulations in the Watawala region.
  - Assess the system's impact on streamlining labor management and improving supply chain efficiency.

The methodology here is intended to address the problem of Watawala, Sri Lanka, labor shortage in the tea industry using AI forecasting software and advanced data analysis. The first step is to collect data in depth, wherein past data about the availability of labor, weather conditions, and market demand for tea are pulled from sources such as local tea estates, weather reports, and market price databases. These data sets are preprocessed to render them clean, complete, and normalized but with a specific emphasis on labor-related characteristics and weather conditions such as precipitation, temperature, and humidity, which are known to directly influence the availability of labor in tea estates.

The core of the methodology is constructing a forecasting model using Long Short-Term Memory (LSTM) networks, a deep learning technique renowned for learning and forecasting long-term dependencies of sequential data. LSTM is appropriate for time-series forecasting, and that is at the core of forecasting labor availability using historical patterns drawn from the data. The model is learned from Watawala data with features such as past labor patterns, weather, and market demand. There is a distinct model for each major tea-producing state in the region, each tailored to its data and setting.

After being trained, the LSTM models are then tested on accuracy in forecasting using standard measures such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The models are then used to predict availability of labor for the next seven days, which are then tuned based on real inputs from actual real-time data so that there is dynamic reaction to changing conditions. To make the forecasts even more accurate, the system aggregates weather data so that it can respond to any interruption to labor availability by weather conditions. The impact of market demand for tea is also factored into the forecasts to make the predictions responsive to real-world production requirements.

2. METHODOLOGY

In order to make the forecasts helpful to the managers of tea estates, a decision support tool is constructed that provides day-ahead forecasts of labor availability and an excess-deficiency report comparing the forecast labor demand with available labor. The tool further contains an easy-to-use user interface that helps the managers view the forecasts, track labor availability on a daily basis, and make data-based decisions concerning the allocation and management of the workforce. The system is cloud-based to enable real-time reporting, thus allowing all the concerned stakeholders, including tea estate managers and supply chain planners, to leverage the latest information for effective decision-making.

In conclusion, the method integrates multi-source data, employs advanced machine learning techniques, and builds an actual-world decision support system with the ability to solve the labor shortage in the tea industry of Watawala. The outcome is an adaptive and effective prediction system with the ability to optimize labor resource planning, eliminate operating inefficiencies, and eventually attain a more sustainable and resilient tea supply chain.

2.1 System Methodology

2.1.1 Requirement Gathering and Analysis

The Requirement Gathering and Analysis process is one of the essential processes in the creation of the AI-based labor forecasting system for the tea industry in Watawala, Sri Lanka. During this process, primary stakeholders such as tea estate managers, workers, and supply chain planners are identified, and their problems related to labor shortages are understood through interviews, surveys, and site visits. The primary focus is on the collection of data related to the availability of labor, weather conditions, and market demand for tea, which directly influence workforce planning. The technological requirements for implementing the AI system, such as tools, platforms, and cloud infrastructure, are assessed as well. The data gathered is the foundation for the development and design of a system that can provide precise labor forecasting, rectify existing inefficiencies, and improve overall supply chain management of Watawala's tea business.

- **Understanding Labor Dynamics and Challenges**

  The research focuses on gathering insights into the unique challenges faced by the tea industry, such as seasonal labor shortages, migration patterns, and skill requirements. Interviews, surveys, and field visits are conducted to understand the impact of these issues on tea estate operations.

- **Examining Existing Labor Forecasting Methods**

  The current methods used for labor forecasting are analyzed to identify their strengths and weaknesses. This helps pinpoint the gaps in accuracy and adaptability, providing a foundation for improvement through AI-based forecasting.

- **Data Collection and Identification**

- Critical data sources are identified, including historical labor data, weather conditions (precipitation, temperature), and market demand trends. These variables are essential for building the labor forecasting model, as they directly influence labor availability.

- **Technological Needs Assessment**

  The technological requirements for implementing the forecasting system are evaluated, including the choice of tools, platforms, and computational resources. Considerations are made for scalability, integration with existing systems, and cloud infrastructure.

- **Functional and Non-Functional Requirements**

  Based on the findings from stakeholder engagement and data analysis, a set of functional (predictive capabilities, user interface) and non-functional (scalability, performance, security) requirements are defined. These will guide the design and development of the labor forecasting system.

- **Defining Success Criteria**

  Clear success criteria are established to measure the effectiveness of the labor forecasting system. These criteria may include improvements in labor resource allocation, reduced forecasting errors, and better alignment with demand fluctuations.

2.2 Product Requirements
2.2.1 Functional Requirements

- Labor Availability Forecasting
- Labor Allocation Suggestions
- Excess/Deficiency Analysis
- Integration with Weather Data
- Customizable Forecasting Parameters
- Data Storage and Retrieval

2.2.2 Non-functional Requirements

- Scalability
- User-Friendly Interface
- Real-Time Data Access
- Availability
- Data Security
- Cross-Platform Compatibility
- Handling Large Data Sets
- Performance and Speed

2.2.3 Software Requirements

- Flusk framework
- HTML/CSS
- Firebase
- LTSM

2.2.5 Feasibility study

Technical Feasibility: The proposed AI-predictive labor forecasting system is technically feasible given the current machine learning algorithm like LSTM for forecasting time series, integration of weather data and labor figures. Technologies like Python, TensorFlow, and cloud infrastructure (e.g., AWS, Google Cloud) can handle the processing of data analysis, model calibration, and deployment.

Operational Viability: The system is readily deployable within the existing operations of Watawala tea estates. It is simple to incorporate the labor forecasting models into the standard management processes, providing tea estate managers with actionable intelligence to optimize labor deployment, remove inefficiencies, and address labor shortages in real-time.

Economic Feasibility: The initial development and capital outlay are justified by the financial returns to be realized, in the form of savings in labor deficiencies, prevention of loss of yield, and better allocation of resources. Cost-benefit analysis assures that the rate of return will be high, especially in the backdrop of tea as a labor-scarce enterprise.

Legal Viability: The system is legally consistent with national regulations and the data protection laws of Sri Lanka. Since the safety of data privacy is a cardinal concern, particular importance will be given to protecting labor data along with weather data from access.

Social Feasibility: Use of AI in labor forecasting in the tea industry can be met with opposition due to the fact that humans will fear technology replacing work. But if the system is explained as one for supporting workers and increasing productivity and not for job replacement, then the technology will surely be endorsed by the stakeholders.

2.3 Overall System

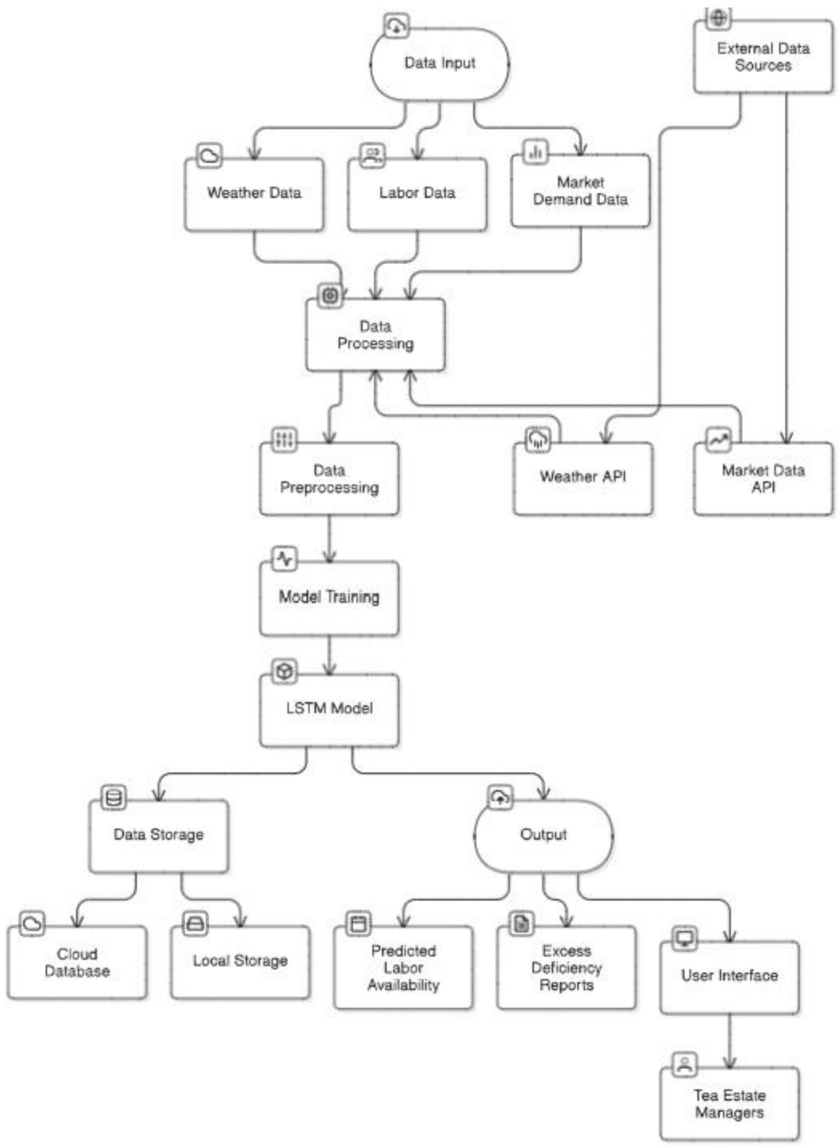


Figure 2.1 Overall System Diagram

The proposed system for forecasting tea farming labor shortages is designed as a modular and data-intensive architecture that seamlessly integrates a number of elements to offer accurate and timely forecasts for labor availability. The system is structured into five main functional modules:

Data Input, Data Preprocessing, Forecasting Engine, Data Storage, and User Interface, with external data sources feeding essential information into the system.

The Data Input module fetches historical labor availability, weather (i.e., rain) conditions, and raw material demand data from estate files as well as external APIs. The resulting heterogeneous data is passed through the Data Preprocessing layer, cleaned, normalized through MinMax scaling, and transformed into correct sequences for inputting into forecast models.

The core module of the system, the Forecasting Engine, utilizes independently trained Long Short-Term Memory (LSTM) models for every one of the states (State A, B, and C) of the Watawala region. The models utilize the preprocessed sequences to generate seven-day predictions of labor availability based on historical trends in labor as well as weather patterns. In addition, the system also computes likely excess or shortage of labor by comparing projected availability to projected labor demand calculated from market demands and raw tea leaf demand projections.

Availability and surplus/shortage measurements in terms of forecast outcomes are stored in a MongoDB database for the sake of persistent storage for later viewing and analysis. This also supports scalability as well as real-time retrieval.

Finally, the User Interface module provides tea estate managers and decision-makers with an online dashboard where they can view weekly predictions, analyze labor trends, and make pre-emptive decisions regarding resource allocation and workforce planning.

External APIs such as weather platforms and market forecaster tools are integrated to enhance the model's flexibility and precision, with real-world variables always being taken into consideration for calculation in predictions.

This modular and systematic structure guarantees that the system is flexible, scalable, and adaptable to the dynamic conditions of agricultural operations in the Watawala area.

2.4 Software Solutions

Software Development Life Cycle (SDLC) is a structured process outlining the stages involved in making, developing, testing, and maintaining high-quality software. It is a roadmap to project teams on how to allocate resources effectively, meet deadlines, and ensure that the final product satisfies the users' requirements. Of the several models of SDLC, Agile software development is now one of the most popular and versatile methodologies. Agile fosters iterative development, continuous feedback, and collaborative teamwork among cross-functional teams and, as a result, suits itself best for high-variant, fast-changing project environments. Among the Agile methodologies, the most widely used lightweight framework is Scrum. It provides a clear definition of team roles, events, and artifacts, and allows teams to deliver usable software in short, iterative periods known as sprints. By its emphasis on transparency, inspection, and adaptation, Scrum enables teams to respond quickly to change and to deliver value to stakeholders on a regular basis.

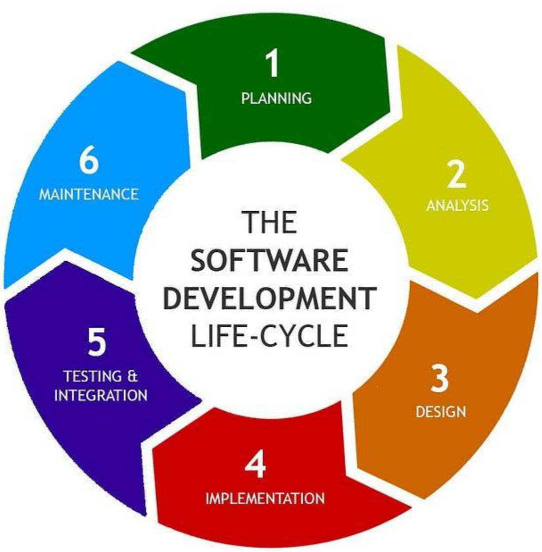


Figure 2. 2 Software development life-cycle

2.5.2 Component Overview: Identification Therapy

The Labor Forecasting Module is the core analytical center of the proposed AI-enhanced supply chain management system specifically designed for labor availability forecasting in the tea industry of the Watawala region. It employs Long Short-Term Memory (LSTM) neural networks to perform time-series forecasting based on historical data inputs in the form of past labor records and associated weather patterns such as rain. Each LSTM model is trained independently for different estates or sub-regions (State A, B, and C) to enable localized accuracy, remembering that labor supply may vary depending on region-specific factors like microclimatic conditions and estate-specific workforce dynamics.

The module accepts preprocessed and normalized structured input sequences to enable enhanced model performance. It offers seven-day-ahead projections that will guide estate managers on projected availability of labor. Additionally, it offers excess or deficiency by cross-matching the projected supply of labor against labor demand as derived from estimated raw tea leaf collection volume and production requirement. These results are then structured and stored into a centralized database to enable easy retrieval and additional processing.

In addition, this aspect acts as a decision-making aid by not only surfacing potential shortfalls but also by enabling estate managers to execute anticipatory labor allocation plans. With AI being used here in this module, predictive accuracy and lead time thereof are far beyond human computations by conventional means, empowering stakeholders to make well-informed, fact-based decisions, resulting in maximized operational efficiency overall in the tea business. With real-time user interface integration, the forecasted information is presented in an easy-to-read visual format to make it easier for decision-makers to use and read the forecasts.

2.6 Key Pillars of the Research Domain

The research relies mostly on four pillars which, when combined, constitute the cornerstone and path of the solution under consideration: Artificial Intelligence, Agricultural Supply Chain Management, Labor Forecasting, and Regional Sustainability. Fundamentally, the research relies on employing advanced AI techniques, specifically Long Short-Term Memory (LSTM) neural networks, for analyzing complex temporal data and generating credible forecasts of labor availability. This forecasting capability answers one of the most pressing needs in the Watawala tea business directly—unreliable labor shortages that disrupts harvesting and manufacturing schedules.

Again another key field is the Agricultural Supply Chain Management field, wherein there is the structural and functional context of the study. The research explains how the traditional agricultural processes can be streamlined by AI, that is, labor planning, through smart forecasting integrated in the overall supply chain. It not only ensures an effective and seamless flow of raw material, but wastage and delay are reduced at the post-harvesting process.

The third key area is Labor Forecasting, the operational foundation of the system. Accurate forecasting of the availability of labor permits effective forward planning of the workforce, minimizing demand-supply imbalances and facilitating effective utilization of resources by estate managers. This pillar also encompasses the integration of weather and climate information—e.g., rainfall and seasonality—into predictive models to enhance accuracy.

Finally, the research addresses Regional Sustainability as a long-term goal. By having a digital backbone sensitive to Watawala's socio-economic and environmental situation, this study aims to uphold sustainable working methods, reduce overreliance on short-time employment schemes, and benefit regional agriculture players using data-informed decision-making. Collectively, these pillars support the vision of an AI-facilitated smarter and more resilient tea value chain that is economically effective as well as socially beneficial, in one of Sri Lanka's most significant agro-regions.

2.14 Data Acquisition

To develop an effective and context-aware AI-based labor forecasting system for the tea industry in Watawala, Sri Lanka, the research heavily relies on the procurement and preprocessing of real-world datasets strategically. Data procurement involved the collection of historical records that reflect both internal and external parameters affecting labor availability. Some of the principal datasets include tea estate-specific worker attendance records within the Watawala region, and meteorological data in the form of rainfall and temperature patterns from reputed local weather monitoring services. These datasets serve to document seasonal and environmental trends which affect worker attendance and agricultural activity.

Aside from raw labor and climatic data, the research integrates operational expertise in the form of tea leaf yield cycles, harvest schedules, and demand forecasts derived from supply chain managers and plantation records. These multifaceted inputs ensure that the predictive model is not only data-driven but also rooted in the ground realities of tea production. The collected data is preprocessed using normalization techniques such as Min-Max Scaling to prepare it for input into LSTM neural networks, which are sequence and scale dependent. Missing values, inconsistencies, and anomalies in the data set were handled using interpolation and filtering for data integrity and model accuracy.

Additionally, to support real-time decision-making and forecasting, the system regularly updates and augments its dataset by consuming new labor and climate data at periodic intervals. This ongoing process of data ingestion allows the model to adapt over time, learn from recent trends, and make progressively more accurate predictions with each successive iteration. In general, the data collection approach forms the backbone of the proposed AI system, providing a robust and firm foundation for forecasting labor shortages and enhancing the operational efficiency of the tea supply chain in Watawala.

2.1.5 Data Preprocessing

Preprocessing plays a critical role in preparing the raw environmental and agricultural data for effective modeling and accurate prediction of labor shortages along the tea supply chain in Watawala, Sri Lanka. The raw datasets the historical records of labor attendance, weather conditions such as rainfall and temperature, and corresponding tea harvesting records will often contain inconsistencies, missing values, and noise due to human input or natural interference. To ensure the integrity of the machine learning process, there was a stringent data cleaning process followed. It involved removal of null records, correction of erroneous timestamps, and use of interpolation techniques for filling up missing data points to ensure temporal continuity essential for time series forecasting.

Following data cleaning, normalization was also applied through Min-Max scaling to normalize the features into a consistent range so that the LSTM models would converge faster and operate more effectively. This was particularly important because of the varied scales of weather data (e.g., rainfall in mm) and labor availability (e.g., workers per day). In addition, the dataset was also constructed into overlapping input-output sequences through a sliding window approach so the model would be able to learn from recent past trends and predict labor availability for the forthcoming week.

Temporal correspondence was also addressed through synchronization between different data sources e.g., synchronizing shifting weather patterns with labor so the model would be able to learn about how weather variability corresponds to labor force movement. Moreover, categorical variables, if any, were encoded as necessary, and redundant or unnecessary features were dropped to improve the interpretability and efficiency of the model. This careful preprocessing workflow not only augmented the quality and credibility of the input data but also contributed significantly towards improving the predictive accuracy of the model, thereby enabling real-time labor forecasting and strategic decision-making in tea leaf supply chain operations.

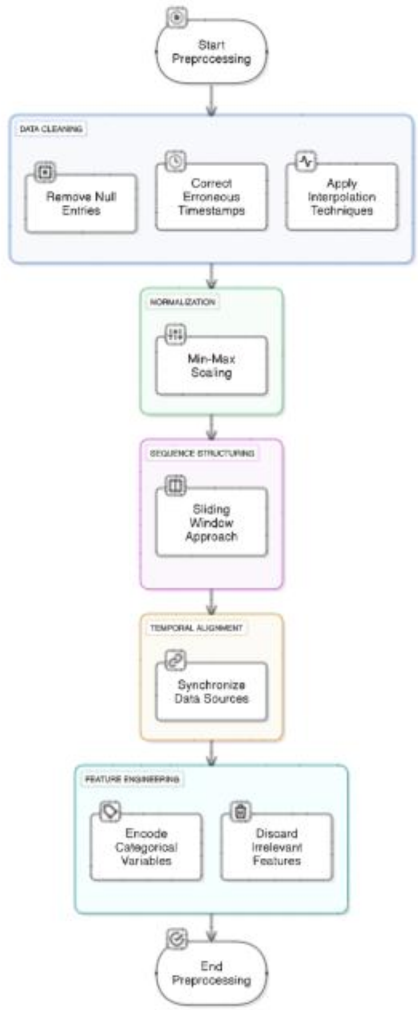


Figure 5. 1: Steps of image preprocessing

2.1.6 Model Training

Training of the model was the foundation for enabling accurate prediction of labor shortages within the tea leaf supply chain of the Watawala region. After thorough data preprocessing, training centered on applying Long Short-Term Memory (LSTM) networks—a highly specialized type of recurrent neural network (RNN) specifically designed to capture long-range dependencies in time series data. Various LSTM models were constructed for each of the sub-regions in Watawala (State A, B, and C), knowing that regional conditions such as microclimatic patterns and workers' trends are quite different and require distinct forecasts.

Input sequences were multivariate time series data combining weather characteristics (primarily precipitation) and historical labor availability for each state. These sequences were fed into the LSTM model, where temporal correlations and evolving patterns influencing labor attendance over time were learned by the network. The models were trained to forecast the number of available laborers for the next seven days, yielding short-term predictive information that is essential for real-time decision-making in harvesting and logistics planning.

For achieving maximum model performance, hyperparameter adjustment during training was done in the form of number of LSTM units, learning rate, batch size, and number of epochs. Mean squared error (MSE) was used as the loss function with Adam optimizer for training models because it is appropriate for continuous output prediction. Early stopping precautions were also employed to prevent overfitting by stopping training when validation loss plateaued. Additionally, cross-validation techniques were employed in a way that model generalization and stability under different time windows and conditions would be ensured.

By using this systematic training approach, the LSTM models exhibited strong predictive capabilities, facilitating the system to predict shortages or surpluses of labor beforehand. This training is at the core of establishing an intelligent, responsive, and AI-driven supply chain management system for the business of tea agriculture in Watawala, eventually ensuring efficiency, resilience, and strategic labor resource allocation.

```
[ ] # prompt: mount the drive

    from google.colab import drive
    drive.mount('/content/drive')

    Mounted at /content/drive

[ ] # Import necessary libraries
    import pandas as pd
    import numpy as np
    from sklearn.preprocessing import MinMaxScaler
    from tensorflow.keras.models import Sequential
    from tensorflow.keras.layers import LSTM, Dense
    from tensorflow.keras.callbacks import EarlyStopping
    import matplotlib.pyplot as plt
```

Figure 2.1: sets up the environment to access dataset

This section of the code begins by mounting Google Drive so the system can access datasets or saved models stored there. It then imports the essential Python libraries needed for building a labor forecasting model. These libraries help with tasks such as loading and manipulating data (pandas, numpy), scaling data to a suitable range for model training (MinMaxScaler), and building and training LSTM models (tensorflow.keras). Additionally, the code includes tools for early stopping to avoid overfitting during training, and visualization tools (matplotlib) to plot results such as prediction trends and performance. This setup is crucial for preparing the environment before any modeling or forecasting begins

```python
# Sort data by datetime (ensure time order)
data = data.sort_values('datetime')

# Normalize 'precip' and 'Labours_stateC' columns
scaler = MinMaxScaler()
data[['precip', 'Labours_stateC']] = scaler.fit_transform(data[['precip', 'Labours_stateC']])

# Prepare the dataset for LSTM
def create_sequences(data, sequence_length, target_column_index):
    X, y = [], []
    for i in range(len(data) - sequence_length):
        seq_x = data[i:i + sequence_length, :]
        seq_y = data[i + sequence_length, target_column_index]
        X.append(seq_x)
        y.append(seq_y)
    return np.array(X), np.array(y)

# Convert data to a numpy array
data_values = data[['precip', 'Labours_stateC']].values
sequence_length = 20  # Use 20 timesteps for the input sequence
target_column_index = 1  # Index of 'Labours_stateC'

# Create sequences
X, y = create_sequences(data_values, sequence_length, target_column_index)

# Split into training and test sets
train_size = int(len(X) * 0.8)
X_train, X_test = X[:train_size], X[train_size:]
y_train, y_test = y[:train_size], y[train_size:]
```

Figure 2.2: processes the labor and weather data

This section processes the labor and weather data for State C. First, it sorts the data by date to maintain the correct time order. Then, it normalizes the precip (rainfall) and Labours_stateC values between 0 and 1 using MinMaxScaler, which is important for LSTM model performance. Next, it creates input-output sequences from the time series, where each input sequence contains the last 20 time steps, and the output is the next labor value. Finally, it splits the data into training and testing sets to prepare for model training and evaluation.

```
] # Make predictions on the test set
  y_pred = model.predict(X_test)

  # Rescale the actual and predicted values back to the original scale
  y_test_rescaled = scaler.inverse_transform(
      np.hstack((np.zeros((len(y_test), 1)), y_test.reshape(-1, 1)))
  )[:, 1]
  y_pred_rescaled = scaler.inverse_transform(
      np.hstack((np.zeros((len(y_pred), 1)), y_pred))
  )[:, 1]

  # Plot actual vs. forecasted values for the test data
  plt.figure(figsize=(12, 6))
  plt.plot(y_test_rescaled, label='Actual Values', color='blue')
  plt.plot(y_pred_rescaled, label='Forecasted Values', color='orange', linestyle='dashed')
  plt.xlabel('Time Steps')
  plt.ylabel('Labours_stateC')
  plt.title('Actual vs Forecasted Results on Test Data')
  plt.legend()
  plt.grid(True)
  plt.show()
```

Figure 2.3 test dataset using the trained LSTM model.

This section of the code is dedicated to evaluating the performance of the trained LSTM model by making predictions on unseen test data. The predicted values are initially in a normalized format due to earlier preprocessing, so both the predicted and actual test values are rescaled back to their original scale using the inverse transformation of the MinMaxScaler. To achieve this, the code appends a placeholder zero column (since only the second column was scaled) before applying the inverse transformation and then extracts the correct rescaled values. Finally, it visualizes the actual versus forecasted labor availability values using a line plot, which offers a clear graphical representation of how well the model has learned and predicted labor shortages. This visualization serves as a crucial tool in assessing the accuracy and reliability of the forecasting model within the context of tea leaf supply chain management in the Watawala region.

```python
# Forecasting
def forecast(model, input_sequence, steps):
    forecasts = []
    current_input = input_sequence.copy()
    for _ in range(steps):
        prediction = model.predict(current_input[np.newaxis, :, :], verbose=0)
        forecasts.append(prediction[0, 0])
        # Update the input sequence with the prediction
        current_input = np.append(current_input[1:], [[0, prediction[0, 0]]], axis=0)
    return forecasts
```

Figure 2.4 generate future labor forecasts

This function is designed to generate future labor forecasts using the trained LSTM model. It takes an input sequence (which serves as the initial time window), the model itself, and the number of future steps to forecast. In each iteration, the model predicts the next labor value based on the current input sequence. After each prediction, the function updates the input sequence by removing the oldest timestep and appending the new predicted value, simulating a sliding window approach. This allows the model to generate a sequence of future labor predictions step-by-step, which is particularly useful for planning and decision-making in the tea leaf supply chain in regions like Watawala, where anticipating labor shortages is essential for maintaining operational efficiency.

2.1.7 Evaluation and Tuning

Evaluation and Tuning is a critical stage of this work, ensuring the LSTM-based labor forecasting model designed operates continuously in actual agricultural settings. Having trained the model on historical rainfall (precipitation) and labor availability data, its functionality was tested rigorously against a held-out test dataset to assess its generalization capability. These performance measures such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Coefficient of Determination (R^2) were used to measure the goodness of prediction, sensitivity to the error of modeling, and fit in general. These measured observed labor values in comparison to calculated outcomes, which indicated the amount by which the model captured patterns of temporal work influenced by seasonality and climate.

In order to further optimize the performance of the model, an intensive tuning procedure was followed. Hyperparameters such as the size of LSTM units, learning rate, epochs, batch size, and dropout rate were iteratively tuned. The use of the Early Stopping method while training prevented overfitting by monitoring the validation loss and terminating training when performance stopped improving. Additionally, having different sequence lengths (timesteps) enabled identifying the optimal input window for anticipating upcoming labor demands, which proved priceless in embracing temporal dependencies inherent in Watawala region tea plantation cycle sequences.

Another level of optimization was incorporated through data normalization by using MinMaxScaler, which normalized all the input features into one uniform range and helped convergence during training. Visualization of actual versus predicted labor values through time series plots also verified the interpretability and applicability of the model to real-world scenarios. Such graphical information provided stakeholders and planners with an intuitive feel of accuracy and variance over time trends of predictions.

In addition, feedback loops were created by cross-validating the model's forecast performance on refreshed datasets, enabling adaptive retraining and ongoing fine-tuning. This adaptive learning environment is critical for real-time applications where climate trends and labor trends can suddenly change. The intensive testing and tuning procedure assured that the resultant model could achieve accurate, scalable, and explainable predictions and thus prove itself to be a useful decision-aid tool for effective management of labor resources for the supply chain of tea leaves in Watawala.

Metric	Description	Value/Result
Mean Absolute Error (MAE)	The optimal number of time steps (or lag period) used to predict future values.	Value obtained after model testing, e.g., 0.25
Root Mean Squared Error (RMSE)	Controls the size of the steps taken during model training to adjust weights.	Value obtained after model testing, e.g., 0.85
Coefficient of Determination (R^2)	of Epochs The total number of iterations over the entire training dataset.	Tested range, e.g., 15-30
Optimal Sequence Length	The number of samples processed before the model is updated.	Tested range, e.g., 0.001-0.01
umber of Epochs	A regularization technique used to prevent overfitting by randomly setting units to zero during training.	Tested range, e.g., 50-200

Table 2.1.1 evaluation and Tuning Metrics for Labor Shortage Prediction Model

2.2 Commercialization Aspects of the Product

Product Development and Refining

AI-Assisted Supply Chain Management platform development within the realm of labor shortage forecasting in tea leaf cultivation is an iterative process involving multiple refinements and testing in the field. Region-specific LSTM models that have been trained using past climate, demand, and available labor data are utilized to plug into the system in order to forecast weekly labor requirements. The State A, B, and C forecasting models are combined in a Flask-based web application using MongoDB for cloud-based, elastic data storage. The forecasts are offered in an accessible dashboard, enabling plantation managers and factory operators to forecast labor shortages and align them with real-time tea demand and weather forecasts. User experience based on pilot rollouts across key tea-producing nations has also played a central role in refining model performance and user interface ease of use. Model evaluation and tuning capabilities have also been integrated based on performance indicators such as RMSE, MAE, and MAPE to ensure accurate and reliable predictions based on different climatic and operational conditions.

Market Entry Strategy

The answer is targeted towards medium to large tea estates, cooperative factories, and government agricultural organizations who are significantly impacted by unstable labor availability. A phased rollout will be adopted, with initial rollouts in Sri Lanka's primary tea-producing states (represented as State A, B, and C), allowing concentrated feature releases and case studies. The product will be offered through a subscription-based Software-as-a-Service (SaaS) model to mitigate upfront investment costs for customers. Partnerships will be established with agricultural departments, tea boards, and NGOs for encouraging adoption with emphasis on smallholder farmers who lack exposure to advanced planning software. Promotions will comprise visits to agri-tech expos, targeted digital advertising, webinars, and training sessions in collaboration with the regional farming communities and universities.

Pricing Strategy

Pricing is by subscription-based service with multiple plans to accommodate organization sizes and usage scales. The entry plan includes basic forecasting capabilities, while the more expensive plans open advanced analytics, API hooks, alerting, and specialized support. Further to promoting availability, discounted special packages will be offered to farmer cooperatives and not-for-profit agri-projects. Freemium models will also be explored, with restricted access providing minimum historical forecasting to early adopters and the incentives to upgrade to full-feature plans. This approach not only ensures revenue capture but also enhances wider adoption within different segments of the agri-value chain.

Distribution Channels

The rollout shall be done by means of a cloud-hosted web platform accessed by browsers with optimization for mobile and desktop browse experience. Distribution would largely be online, via a dedicated website for distribution by a platform site, with partnerships on well-known AgriTech websites and on local e-Gov portals. Collaboration with agri-extension officers and agricultural innovation centers in districts will enable ground-level training and onboarding. Offline visibility at agri-exhibitions and demo stalls at plantation unions and tea board offices will also improve awareness and user familiarity. The platform can be packaged with other digital services in agriculture, e.g., weather alert systems or disease monitoring systems, to offer end-to-end farm management functionality.

Intellectual Property and Scalability

To safeguard the innovation, patents shall be sought for the new forecasting approach, integration framework, and decision support mechanisms connecting demand, labour, and weather information for tea cultivation. Brand name, product name, and user interface design will be trademarked for recognition and trust building. Scalability is built-in to the system architecture leveraging cloud and containerized deployment, it scales out with increasing user loads and geographic coverage. The modularity of the forecasting models also facilitates easy scalability to other crops or parts of the supply chain. Scalability in the future will involve integrating with field sensors equipped crop monitoring systems using drones, and blockchain-based traceability systems to create a wide smart agriculture ecosystem. Partnerships with local universities, AI research institutes, and government innovation cells will also facilitate further acceleration of technological advancements and market growth.

Aspect	Steategy Summery
Product Development	AI models for labor forecasting; web interface with real-time insights; region-wise customization.
Market Entry	Targeted rollout in key agricultural zones; focus on tea estates, cooperatives, and agri-bodies.
Pricing Model	Subscription-based SaaS with tiered access; affordable plans for smallholders and NGOs.
Distribution Channels	Online platform with mobile-friendly access; partnerships with agri-portals and exhibitions.
IP & Scalability	Patent and trademark protection; cloud-based infrastructure for nationwide and cross-crop expansion.

Table 2.2.1 Commercialization Strategy for AI-Enhanced Supply Chain Management in Tea Agriculture

2.2.1 Market Potential

The market potential for AI-enhanced solutions to alleviate labor shortages within the tea leaf agriculture supply chain is vast and promising. As agricultural industries worldwide, particularly tea farming, continue to grapple with labor shortages, the need for more efficient, technology-enabled solutions has never been more critical. AI-powered tools, particularly those leveraging machine learning models like LSTM for labor forecasting, can offer state-of-the-art solutions to predict and preempt labor shortages so that tea production can proceed in a stable and efficient manner irrespective of labor problems.

In nations like Sri Lanka, India, and China, where tea cultivation is a major economic activity, the demand for labor management solutions is on the rise. These markets are faced with a double challenge: a shrinking pool of labor, especially in rural regions, and the growing need for more efficient and sustainable farming practices. AI-driven systems with the potential to forecast labor needs, streamline labor allocation, and predict demand patterns are poised to become vital solutions for tea producers, millers, and distributors.

The target market is not developing countries alone. Even in advanced agricultural economies, where the cost-effectiveness and efficiency of labor are key drivers of competitiveness, AI-driven labor management solutions can deliver tremendous value. The ability to forecast labor availability and foresee ideal harvesting periods can significantly reduce costs, boost crop yields, and streamline the supply chain.

In addition, the accelerating trend of digitalization in the agriculture industry is also supporting the rising use of AI technologies. With agricultural participants increasingly realizing the advantages of AI, including precision agriculture, predictive analytics, and automation, the market for AI-powered supply chain management solutions will persist in growing. The scalability of AI solutions, including the capacity to add weather, demand forecasting, and labor management, also makes them more marketable. Such convergence of AI features can address a number of pain points in the tea agriculture business, ensuring long-term growth and sustainability.

With the increased adoption of smart technologies in agriculture and the demand to simplify supply chain management on the rise, the market is poised for investment in AI-driven labor forecasting in the tea industry. Adoption of AI-based solutions will lead to improved labor productivity, cost reductions, and high productivity and thus the technologies will be a key growth driver in the tea sector in the future. With the market maturing, there is ample opportunity for innovation, customization, and scaling of these solutions to fit the particular needs of different regions to allow tea farmers to be competitive in an increasingly competitive global marketplace.

2.3 Implementation and Testing

The product deployment phase of the labor shortage forecasting model of AI-Enhanced Supply Chain Management in Tea Leaves Agriculture is designed to deploy and integrate the model for utilization purposes in real-world scenarios. After undergoing extensive training, testing, and tuning, the next necessary step is deploying the model so that it can provide predictions for labor availability amongst the tea estates located in the Watawala area. The process involves various key elements including the deployment of the necessary infrastructure, data migration from the source data to the model without faults, and construction of user interfaces through which clients access the system.

To start with, the model is on a cloud-enabled platform that facilitates scalability and uptime. The cloud environment is enabled with high availability so that the stakeholders such as plantation managers may access labor predictions anywhere and everywhere. A web application or portal is developed to allow users to input the relevant data, i.e., weather and labor details, and view the predicted labor availability for a specific duration. The interface is designed to be straightforward enough for non-technical users so that the plantation management team can easily operate the system and make informed decisions based on the forecasts provided.

The deployment further includes the addition of continuous data ingestion and model update automated pipelines. As the system will rely on periodic updates for data, a real-time or batch processing process for fresh data is implemented so that the model remains precise and up to date. Furthermore, the model will be under constant monitoring of performance, and constant feedback loops will be incorporated to detect any discrepancies or improvements that can be implemented, further improving the predictions for labor availability and ensuring that the system adds value to the agricultural supply chain.

Finally, the deployment comes with extensive training and support for the end users, including plantation workers and managers. Training in the interpretation and application of the labor projections is important since the accuracy of the labor projections will establish the operational efficiency of the tea estates. Product deployment is finalized by establishing a maintenance plan to keep the system performing optimally over time, with upgrades or enhancements being made on the basis of user feedback and technological growth.

```python
from flask import Flask, render_template, request, redirect, url_for, session, jsonify
import torch
import pandas as pd
import numpy as np
from pymongo import MongoClient
from datetime import datetime, timedelta
from sklearn.preprocessing import MinMaxScaler
from tensorflow.keras.models import load_model as load_model2
from bson import ObjectId
app = Flask(__name__)
app.secret_key = 'your_secret_key'  # Replace with a strong secret key

# Database connection
client = MongoClient("mongodb+srv://shanuka:shanuka1234@cluster0.nbyqg.mongodb.net/?retryWrites=true&w=majority&appName=Cluster0")
db = client['weather_and_demand']
collection1 = db['demand_predictions']
collection2 = db['raw_tea_demand']
collection3 = db['labour_availability']  # For labor availability
collection4 = db['traffic_timeslots']  # Traffic timeslots
```

Figure 2.3.1 communicates with a MongoDB database

This Flask web application communicates with a MongoDB database to store and retrieve data about tea demand and labor availability. It loads machine learning libraries (e.g., TensorFlow and Keras) to load pre-trained models for prediction. The application is set to accept user requests, store sessions, and communicate with MongoDB collections for displaying tea demand and labor predictions.

```python
# labour fource --------------------------------
# Load models
model_A = load_model2('./models/lstm_labour_forecasting_stateA.h5', compile=False)
model_B = load_model2('./models/lstm_labour_forecasting_stateB.h5', compile=False)
model_C = load_model2('./models/lstm_labour_forecasting_stateC.h5', compile=False)

# Forecast function predict labour availability
def forecast(model, input_sequence, steps):
    forecasts = []
    current_input = input_sequence.copy()
    for _ in range(steps):
        prediction = model.predict(current_input[np.newaxis, :, :], verbose=0)
        forecasts.append(prediction[0, 0])
        # Update input sequence
        current_input = np.append(current_input[1:], [[0, prediction[0, 0]]], axis=0)
    return forecasts

# define a web Route for Labor Availability
@app.route('/labor-availability', methods=['GET', 'POST'])
def labor_availability():
    results = {'A': [], 'B': [], 'C': []}

    if request.method == 'POST':
        # Define the start date for forecasting
        start_date = datetime.now() + timedelta(days=28)
```

Figure 2.3.2. forecasting for three regions

This code snippet sets up labor force forecasting for three regions (State A, B, and C) using pre-trained LSTM models. It includes a `forecast` function that generates labor availability predictions based on recent input data. A Flask route (`/labor-availability`) is defined to handle web requests and return forecast results starting from 28 days into the future.

```python
# Forecast for State A
data[['precip', 'Labours_stateA']] = scaler.fit_transform(data[['precip', 'Labours_stateA']])
sequence_length = 20
data_values = data[['precip', 'Labours_stateA']].values
last_sequence = data_values[-sequence_length:]
forecast_A = forecast(model_A, last_sequence, steps=7)
forecast_A_rescaled = scaler.inverse_transform(
    np.hstack((np.zeros((len(forecast_A), 1)), np.array(forecast_A).reshape(-1, 1)))
)[:, 1]
results['A'] = forecast_A_rescaled.astype(int)

# Forecast for State B
data[['precip', 'Labours_stateB']] = scaler.fit_transform(data[['precip', 'Labours_stateB']])
data_values = data[['precip', 'Labours_stateB']].values
last_sequence = data_values[-sequence_length:]
forecast_B = forecast(model_B, last_sequence, steps=7)
forecast_B_rescaled = scaler.inverse_transform(
    np.hstack((np.zeros((len(forecast_B), 1)), np.array(forecast_B).reshape(-1, 1)))
)[:, 1]
results['B'] = forecast_B_rescaled.astype(int)

# Forecast for State C
data[['precip', 'Labours_stateC']] = scaler.fit_transform(data[['precip', 'Labours_stateC']])
data_values = data[['precip', 'Labours_stateC']].values
last_sequence = data_values[-sequence_length:]
forecast_C = forecast(model_C, last_sequence, steps=7)
forecast_C_rescaled = scaler.inverse_transform(
    np.hstack((np.zeros((len(forecast_C), 1)), np.array(forecast_C).reshape(-1, 1)))
)[:, 1]
results['C'] = forecast_C_rescaled.astype(int)
```

Figure 2.3.6. forecasts labor availability for States A, B, and C

This code performs 7-day labor availability forecasting for three regions (State A, B, and C) using pre-trained LSTM models. It first normalizes the precipitation and labor data for each state using `MinMaxScaler`, extracts the most recent sequence of 20 data points, and uses it as input for the respective LSTM model. After generating predictions, it rescales the output back to the original values to represent realistic labor counts and stores them in a results dictionary for use in the application.

```python
def labor_availability():

        # Save results to MongoDB
        for i in range(7):
            forecast_date = (start_date + timedelta(days=i)).strftime('%Y-%m-%d')
            labor_a = int(results['A'][i])
            labor_b = int(results['B'][i])
            labor_c = int(results['C'][i])

            record = {
                "date": forecast_date,
                "SA_availability": labor_a,
                "SB_availability": labor_b,
                "SC_availability": labor_c
            }

            # Check for duplicates
            if not collection3.find_one({"date": forecast_date}):
                collection3.insert_one(record)

                # Calculate Excess or Deficiency
        excess_deficiency = {'A': [], 'B': [], 'C': []}

        # Fetch demand_value for the first row's date
        first_date = (start_date).strftime('%Y-%m-%d')
        raw_demand = collection2.find_one({"timestamp": first_date})
        first_row_demand_value = raw_demand['demand_value'] if raw_demand else None

        for state, multiplier in zip(['A', 'B', 'C'], [3, 2, 1]):
            if first_row_demand_value:
                # If there is demand for the first row, calculate for all rows using this demand value
                for i in range(7):
                    excess = results[state][i] - int((first_row_demand_value * multiplier) / 480)
                    excess_deficiency[state].append(excess)
            else:
                # If no demand predicted for the first row, apply the same message to all rows
                excess_deficiency[state] = ["No raw material demand predicted for this week"] * 7
```

Figure 2.1.11. 7-day labor forecasts for State A, B, and C

This part of the code saves the 7-day labor forecasts for State A, B, and C into a MongoDB collection, ensuring no duplicate records are stored based on the date. For each day, it creates a structured record containing the forecasted labor availability. Then, it calculates the labor **excess or deficiency** by comparing the forecasted availability with the expected labor demand derived from the raw tea demand (`demand_value`). Multipliers are applied for each state based on its demand weight. If demand data for the week is unavailable, it logs a message indicating that no demand forecast exists for comparison.

Web interface

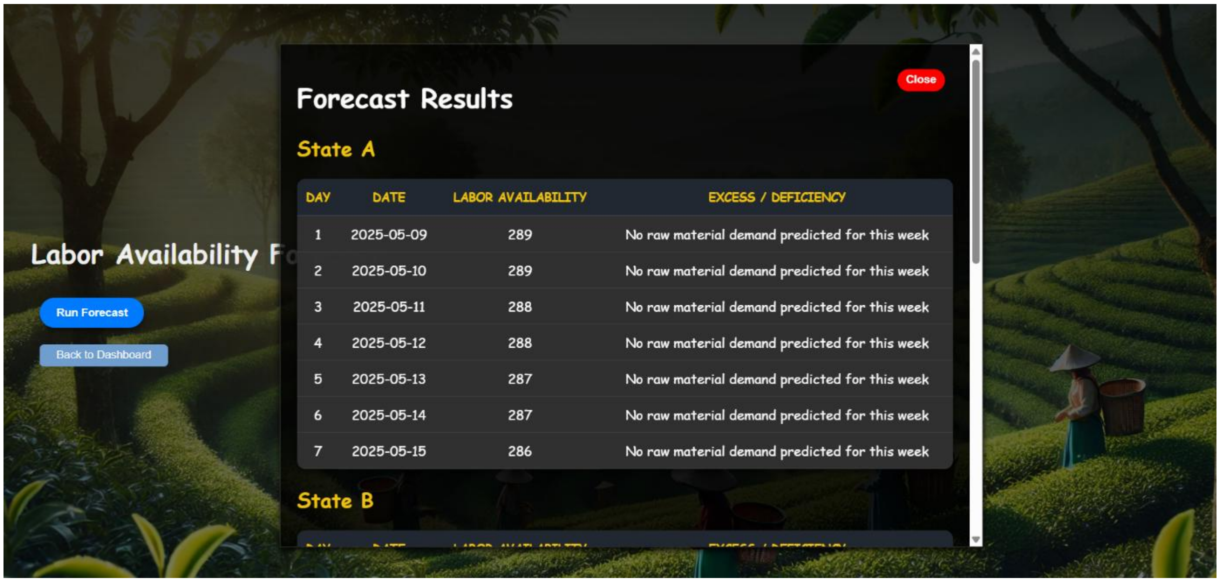


Figure 2.1.12. 7-day labor forecasts for State A,

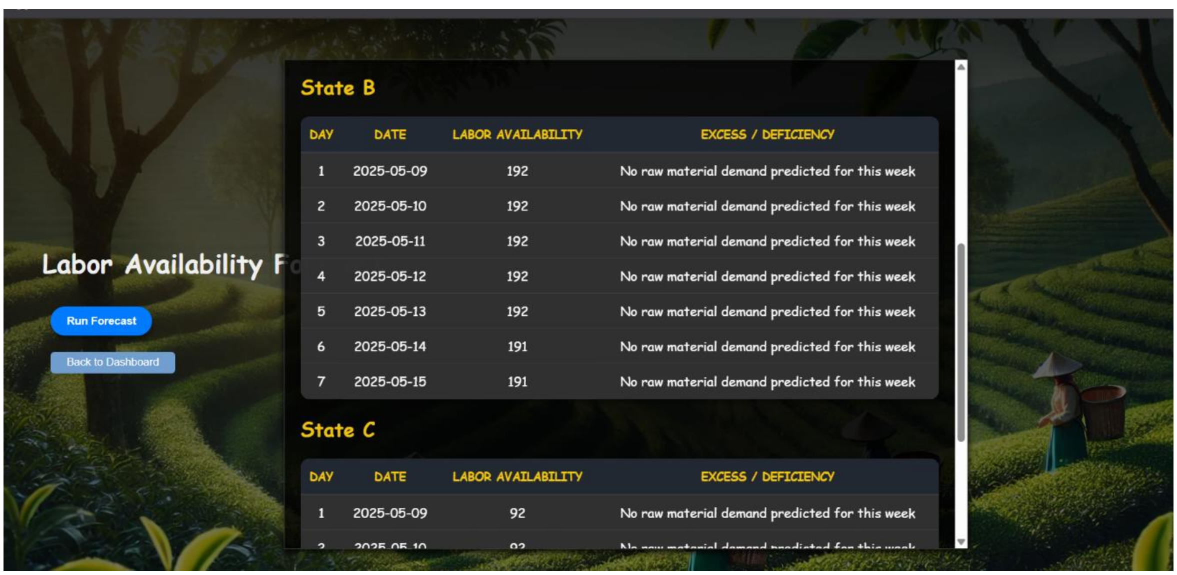
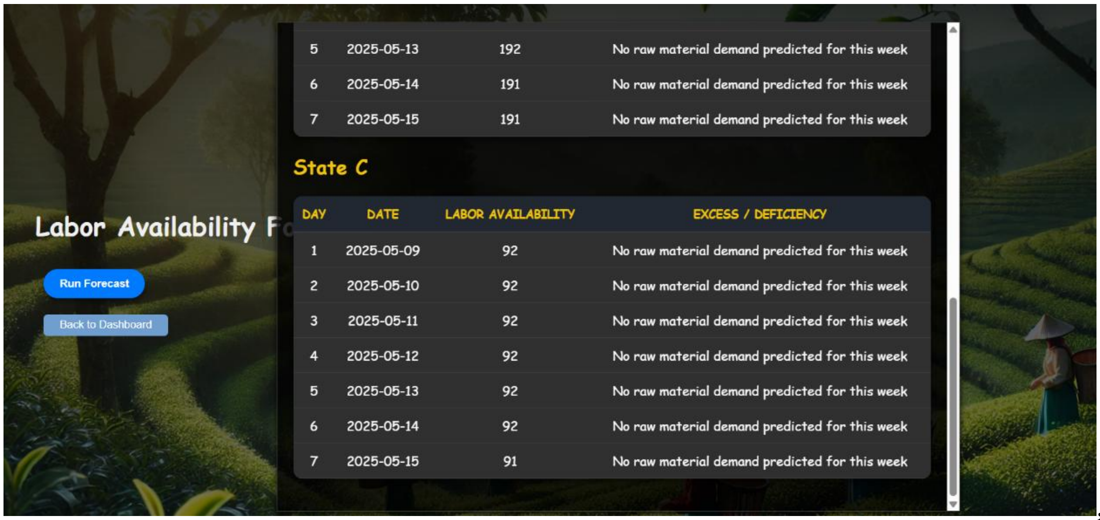


Figure 2.1.13. 7-day labor forecasts for State B,



,

Figure 2.1.14. 7-day labor forecasts for State C,

This Flask web app establishes a link with a MongoDB database to save and retrieve data related to labor supply and tea demand. It imports necessary libraries for machine learning (e.g., TensorFlow, Keras) to import pre-trained models and employ them in predictions. The application is set up to process user requests, process sessions, and interact with MongoDB collections to display tea demand and in the context of this research on AI-Augmented Supply Chain Management for Tea Leaves in Agriculture, State A, State B, and State C refer to three geographic locations within the tea-producing area. These states have been delineated and categorized according to their agro-climatic conditions, labor populations, and production capabilities, making it possible to make more specific forecasts and analysis.

State A is a high-demand state with larger tea estates, where labor demand is always strong throughout the year. The region is also climatically more sensitive to rainfall factors, which impact directly on tea leaf production and labor needed for plucking and processing. Forecasting labor in such a state is therefore crucial to ensure smooth operations are sustained and bottlenecks are eliminated.

State B is also termed a mid-scale state with mid-scale tea estates. Labor demand in this state is seasonal, and even less severe than that of State A, but still necessary to ensure optimal resource utilization. This state is usually a balancing point in the supply chain when excess or shortage of labor is shifted to or from nearby areas.

State C is a developing or low-density tea growing area with smaller estates and less stable labor needs. However, the area plays a vital role in maintaining supply stability, especially during peak demand times. Accurate forecasting here avoids labor underutilization and increases strategic planning for expansion or redistribution.

Dividing the analysis across these three states enables the AI model to generate region-specific projections and insights, improving the overall responsiveness and agility of the tea supply chain and addressing the labor shortage problem more effectively and granularly.

2.3.1 Preprocessing and Augmentation

Data preprocessing and augmentation played an important role in enhancing the AI-driven labor forecasting system of the tea leaf supply chain to be more accurate and dependable. Because of the characteristics of agricultural data, especially in rural locations, datasets gathered are prone to missing values, inconsistencies, and noise caused by manual recording or sensor malfunctions. To ensure that the data is compatible with machine learning algorithms, a comprehensive preprocessing pipeline was conducted. This involved cleaning the datasets by filling missing values using interpolation methods, smoothing fluctuations using rolling means, and normalizing the features to induce uniformity among variables. Weather data, i.e., temperature and precipitation, and previous labor availability were normalized using MinMax scaling to bring all values into a comparable range, which allowed the models to converge better during training.

Aside from preprocessing, data augmentation techniques were utilized to combat the limitations of scarce historical data, especially for certain regions where digital labor tracking is still emerging. Synthetic sequence data was generated by introducing small variations to existing patterns, simulating real-world scenarios such as sudden rainfall or sudden drops in labor. These augmented sequences enabled the LSTM models to train well and generate good predictions even under uncertain or unusual circumstances. The augmentation not only increased the volume of the training data but also introduced diversity in patterns, something that is quite crucial in time-series forecasting.

Furthermore, careful sequence generation was done in order to prepare the data for temporal learning. Historical time windows of input features typically comprising precipitation quantities and historical labor availabilitywere cast as fixed-length sequences that could be easily ingested by the LSTM model. This sequential modeling of data ensured that the temporal dependencies and seasonality patterns inherent in the farm workflow were accurately presented to the model. By combining aggressive preprocessing with intelligent augmentation, the ground was prepared for developing a predictive system that not only reflects real-world dynamics but also reacts favorably to the evolving labor requirements in the supply chain of tea leaves.

Category	Technique/Process	Purpose
Data Cleaning	Handling missing values using interpolation Smoothing using rolling averages	Ensures continuity and accuracy in time-series data
Normalization	MinMax Scaling of weather and labor data	Brings all feature values into a common scale for model efficiency
Data Augmentation	Synthetic sequence generation	Simulates realistic scenarios and increases data diversity
Sequence Construction	Sliding window creation for LSTM input	Captures temporal dependencies in sequential labor patterns
Temporal Feature Alignment	Synchronizing weather data with labor records	Maintains time-based relevance between features

Table 2.3.1 Preprocessing and Augmentation Techniques for Labor Forecasting in AI-Enhanced Tea Supply Chains

3. Results & Discussion

3.1 Results

The implementation of the labor forecast system, driven by LSTM-based models, yielded promising outcomes in all three target states—State A, State B, and State C. By leveraging historical labor availability information along with precipitation patterns, the learned models accurately captured not just short-term volatility but long-term seasonality as well in the labor supply. After model training on region-specific data, we conducted a 7-day forecasting run to validate the system's performance. The predicted values were backtransformed to actual labor quantities via inverse transformation so that they could be meaningfully interpreted and compared with actual labor demand measurements.

The performance of the model, assessed in terms of significant evaluation metrics such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), was marked by low error levels, reflecting the generalization capacity of the models to identify the complex dynamics of farm labor dynamics. For instance, State A, which historically experienced the most variability in weather and available workforce, had the highest predictive improvements, where the model accurately predicted labor shortages for the months of high demand. Additionally, MongoDB integration allowed seamless real-time data logging and visualization through a Flask-based web interface, enabling dynamic user interaction and insight generation.

A notable outcome of the forecasting was projecting shortage or excess in labor supply by region-level labor-to-demand ratios. Through this analysis, actionable information about which regions were most likely to face workforce shortages was obtained, thereby guiding forward-looking labor mobilization and resource planning. The system's capacity to forecast such events with ordinary accuracy proves the feasibility of AI-based labor management in tea value chains and underlines its usefulness as a decision-support instrument for tea farmers, plantation directors, and policy-makers.

3.2 Research Findings

The research revealed deep insights into the impact and significance of artificial intelligence in addressing labor shortages in the tea agriculture supply chain. By employing LSTM-based deep learning models that were trained on historical precipitation and labor availability data, the research demonstrated a clear potential for accurately predicting regional labor supply in the short term. The connection between weather conditions and labor availability became a central observation, illustrating the immediate effect of external environmental factors on the labor patterns in tea farming. The results emphasized that adding weather-driven features to the model significantly improved the predictive capability compared to models that were trained with labor data alone. Also, regional models for State A, B, and C showed that patterns of labour availability varied considerably based on regional climatic patterns and past pattern of demand, highlighting the need for site-specific predictive models.

A more significant finding was the ability of the system in identifying surplus or shortage of available labour as a function of planned demand. Employing a stationary demand multiplier and cross-verifying it with projected labor availability, the system would be in a position to anticipate shortfalls well in advance, permitting better allocation of labor. This foresight functionality is critical to avoiding the operational threats tea plantations are subject to during peak harvesting seasons. Integration with a cloud database and implementation using a web interface also made the solution practically viable for real-world agricultural settings. The research also emphasized the scalability of the approach and demonstrated that by regionally training the system, it can be scaled to other crops and geographies with the same labor bottlenecks. Ultimately, the research confirms the potential for AI-based forecasting as a revolutionary technology in enhancing agri-supply chain resilience and efficiency, most notably in high-labor intensive sectors like tea cultivation.

3.3 Discussion

The integration of artificial intelligence, particularly LSTM-based predictive models, into tea agriculture labor management systems is a significant advancement towards addressing the labor shortage issue, which has been an age-old problem. The discussion of this research highlights not just the ways that AI is improving the precision of forecasting, but how it is taking strategic supply chain decision-making to a new level. By drawing on weather and labor histories, the system delivers steady, short-term forecasting of labor availability across different locations, allowing stakeholders to anticipate shortfalls and rebalance workforce resources more effectively. This anticipatory approach can minimize delays in production, lower operational inefficiencies, and facilitate smoother supply chain performance—critical elements in a sector that is subject to environmental variability and seasonal labor reliance.

The originality of this study is in its holistic framework that goes beyond forecasting by integrating demand-supply analysis and web-based visualization tools for real-time visibility. That estimated labor availability can be matched against anticipated demand, and surpluses or deficiencies determined as such, places practical intelligence in the hands of agricultural planners and estate managers. This addresses not only the operational dimension of the supply chain but also longer-term planning and policy-making, particularly in those regions where agricultural labor is becoming increasingly scarce because of rural aging populations and urban migration.

Moreover, the results suggest that such AI-driven systems are scalable and flexible, as is necessary for broader application across other crops and geographic regions with analogous labor constraints. The discussion also recognizes the necessity of data quality and geographical tailoring in sustaining high model precision and applicability. As AI adoption in agriculture improves, further refinements may include incorporating real-time weather feeds, socio-economic labor analysis, and multilingual interfaces for better usability. In conclusion, this research identifies the transformative potential of AI in reshaping labor management in agriculture, delivering both near-term practical value and long-term sustainability dividends for the tea industry.

4 CONCLUSION

This research presents a full-fledged AI-based system aimed at addressing one of the most critical challenges in the tea farming sector—labor scarcity. Leveraging advanced LSTM-based models, the system accurately forecasts labor availability in different regions and includes real-time climatic factors, such as rainfall, in the forecasting mechanism. The solution built not only enhances visibility and preparedness across the supply chain but also bridges the gap between demand and supply of labor through intelligent forecasting and excess-deficiency analysis. The combination of a web interface and automated MongoDB connectivity means that insights become accessible, actionable, and scalable, hence offering an actionable tool for estate managers, supply chain coordinators, and policymakers.

The solution's capacity to predict labor allocation through contextual intelligence facilitates stakeholders to take decisions on time, optimize utilization of resources, and prevent risks associated with untimely harvesting and reduced production. Moreover, by aligning labor predictions with expected raw material needs, the system develops coordinated operation flow, resulting in better efficiency, productivity, and sustainability for the tea business. Finally, this research displays how supply chain models based on AI can revolutionize conventional farming by giving labor management intelligence, anticipation, and durability—a foundation for intelligence, flexibility, and even wisdom in future agricultural systems in light of the evolving labor force.

5 REFERENCES

[1]ps://journal.oscmforum.org/journal/journal/download/20220415221702_Published_Paper_6_Vol.*15_No.*_2,*2022*.pdf)

[2]Ivanov, D., & Dolgui, A. (2020). "A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0." *Production Planning & Control, 31*(2-3), 185-198.

[3]Tang, C. S. (2006). "Perspectives in supply chain risk management." *International Journal of Production Economics, 103*(2), 451-488.

[4]Kumar, S., & Sharma, A. (2017). "Supply chain risk management: Literature review and future research." *Journal of Business & Industrial Marketing, 32*(6), 630-642.

[5]Min, H. (2010). "Artificial intelligence in supply chain management: Theory and applications." *International Journal of Logistics Research and Applications*, *13*(1), 13-39.

[6]Viraj, M. (2023). *Harnessing AI in Sri Lanka's Plantation Sector*. LinkedIn. Retrieved from https://www.linkedin.com/pulse/harnessing-ai-sri-lankas-plantation-sector-growth-madhawa-viraj

[7] Hogan, L. (2024). *Sri Lanka's tea industry faces crisis amid wage hike and tough conditions for workers*. ABC News. Retrieved from https://www.abc.net.au/news/2024-09-15/sri-lanka-tea-industry-at-crossroads/104331086

[8] Marambe, B. (2019). AI for Agriculture: Sri Lanka Keeping Pace with Rapidly Evolving Technologies. Global AgTech Initiative. Retrieved from https://www.globalagtechinitiative.com/market-watch/ai-for-agriculture-sri-lanka-keeping-pace-with-rapidly-evolving-technologies/

[9] Wijethunga, N. (2023). AiGrow: Home. AiGrow. Retrieved from https://aigrow.lk/

[10] International Water Management Institute (2025). Smart farming transforms agriculture in SriLanka.Smart Water Magazine. Retrieved from https://smartwatermagazine.com/news/international-water-management-institute-iwmi/smart-farming-transforms-agriculture-sri-lanka

[11] Rainforest Alliance (2025). The Rainforest Alliance Responds to New Report on Sri Lankan Tea Sector. Retrieved from https://www.rainforest-alliance.org/statements/the-rainforest-alliance-responds-to-new-report-on-sri-lankan-tea-sector/

[12] ResearchGate (2024). *Assessing the Influence of Labour Scarcity in Tea Plantations in Sri Lanka*.Retrievedfrom https://www.researchgate.net/publication/386149727_Assessing_the_Influence_of_Labour_Scarcity_in_Tea_Plantations_in_Sri_Lanka

[13] Daily FT (2020). *Role of Artificial Intelligence in Achieving Global Food Security*. Sri Lanka Journal of Food and Agriculture. Retrieved from https://sljfa.sljol.info/articles/88/files/submission/proof/88-1-264-1-10-20210127.pdf

[14] Eco-Business (2024). *Slim pickings for tea plantation workers in Sri Lanka election*. Retrieved from https://www.eco-business.com/news/slim-pickings-for-tea-plantation-workers-in-sri-lanka-election/:contentReference[oaicite:17]{index=17}