Journal of Infrastructure, Policy and Development 2025, 9(2), 8966.
https://doi.org/10.24294/jipd8966

Article

Leveraging deep learning for agricultural market forecasting and supply chain optimization in Malaysia

Mirza Nurhaziq Mazatul Affendy, R. Kanesaraj Ramasamy*, Venushini Rajendran

Faculty Computing Informatics, Multimedia University, Cyberjaya 63100, Malaysia
*** Corresponding author:** R. Kanesaraj Ramasamy, kanes87@gmail.com, r.kanesaraj@mmu.edu.my

CITATION

Affendy MNM, Ramasamy RK, Rajendran V. (2025). Leveraging deep learning for agricultural market forecasting and supply chain optimization in Malaysia. Journal of Infrastructure, Policy and Development. 9(2): 8966. https://doi.org/10.24294/jipd8966

ARTICLE INFO

Received: 3 September 2024
Accepted: 22 November 2024
Available online: 10 February 2025

COPYRIGHT

Copyright © 2025 by author(s). *Journal of Infrastructure, Policy and Development* is published by EnPress Publisher, LLC. This work is licensed under the Creative Commons Attribution (CC BY) license. https://creativecommons.org/licenses/by/4.0/

Abstract: The Agriculture Trading Platform (ATP) represents a significant innovation in the realm of agricultural trade in Malaysia. This web-based platform is designed to address the prevalent inefficiencies and lack of transparency in the current agricultural trading environment. By centralizing real-time data on agricultural production, consumption, and pricing, ATP provides a comprehensive dashboard that facilitates data-driven decision-making for all stakeholders in the agricultural supply chain. The platform employs advanced deep learning algorithms, including Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN), to forecast market trends and consumption patterns. These predictive capabilities enable producers to optimize their market strategies, negotiate better prices, and access broader markets, thereby enhancing the overall efficiency and transparency of agricultural trading in Malaysia. The ATP's user-friendly interface and robust analytical tools have the potential to revolutionize the agricultural sector by empowering farmers, reducing reliance on intermediaries, and fostering a more equitable trading environment.

Keywords: agriculture trading platform; deep learning algorithms; market forecasting; data-driven decision making; agricultural supply chain

1. Introduction

1.1. Overview

The agricultural sector plays a pivotal role in Malaysia's economy, contributing significantly to domestic consumption, food security, and export revenues. In 2021, agriculture accounted for 4.1% of Malaysia's GDP, with key exports including palm oil, rubber, and timber generating substantial income. Notably, Malaysia was the world's largest exporter of palm oil, earning RM92.2 billion (USD22.5 billion) in 2021, while rubber and timber exports contributed RM20.8 billion (USD5.1 billion) and RM20.5 billion (USD5.0 billion), respectively. Despite self-sufficiency in staple foods like rice, poultry, and vegetables, Malaysia still relies on imports for livestock and dairy products.

Consumption patterns across Malaysia reveal a dependence on rice as the staple food, followed by poultry and vegetables, with rural areas showing lower consumption of dairy and beef. The Household Income and Expenditure Survey (2018/2019) found that the average monthly household food expenditure was RM545.20 (USD133.30), with rice accounting for 26.4% of the budget, followed by poultry (17.5%) and vegetables (13.4%).

Globally, digital platforms are transforming the agricultural landscape by connecting farmers with consumers, reducing intermediaries, and providing real-time market insights. Successful examples include Alibaba's Yunshang Platform and JD.com's Jingdong Rural E-commerce Platform in China, which enhance market access and profitability for farmers. Platforms like AgNext, Gramophone, and MFarm offer market information, financial services, and real-time pricing, underscoring the potential of technology to revolutionize agriculture. However, challenges remain, such as building trust between buyers and sellers, addressing infrastructure gaps, and ensuring equitable technology access.

To empower producers and enhance market efficiency, agricultural trading platforms must integrate predictive algorithms to anticipate market trends and price fluctuations. They should provide direct access to domestic and international buyers, reducing reliance on intermediaries. Clear data visualizations and transparent transactions are crucial for enabling producers to monitor costs, optimize sales, and receive fair compensation. This study introduces a comprehensive platform designed to address these needs, leveraging advanced technologies to improve decision-making and market accessibility for Malaysia's agricultural stakeholders.

1.2. Project statement

The current agricultural trading environment is characterized by inefficiencies and a lack of transparency. Farmers face challenges in finding buyers for their produce, while buyers struggle to obtain high-quality products at reasonable prices. Furthermore, the absence of a centralized market information system leads to information asymmetry, hindering market participants from making well-informed decisions. The objective of this project is to address these issues by developing a web-based application platform that promotes transparent, efficient, and easily accessible agricultural trading.

1.3. Project objectives

Objective 1: To design a dashboard that forecast the Malaysia main agriculture commodities being produce and where mostly consume and market price trends.

Outcome 1: A centralized web-based dashboard that cater the action for forecasting agriculture commodities market price and forecasting consumption area of specific commodities.

Objective 2: To understand the deep learning algorithm used in order to predict and locate Malaysia agriculture market trends.

Outcome 2: An implementation of LSTM in performing forecasting for market price and CNN for forecasting consumption areas.

Objective 3: To develop and promote better market linkages for farmers which will help farmers access larger markets and negotiate better prices.

Outcome 3: An effective platform created to link connections between farmer and buyer throughout the system.

1.4. Project scopes

The scope of this project is to revolutionize agricultural trading by developing a user-friendly Agriculture Trading Platform (ATP) that is accessible via the web. The platform will facilitate the listing, searching, and trading of agricultural products, while also integrating real-time pricing data, providing detailed product specifications, and improving communication. The focus is on the creation of a prototype that will address the inefficiencies and lack of transparency in the current agricultural trading landscape. This will improve the accessibility of market information and product quality data, as well as connect farmers, producers, distributors, and consumers. All users have access to the Agriculture Trading Platform (ATP).

1.5. Deliverable

1) An intuitive interface that presents up-to-the-minute information on agricultural commodities in Malaysia. The features encompass map visualization, comprehensive overviews of production and consumption, trade movements, and real-time updates on pricing trends.
2) A predictive system employing deep learning algorithms to estimate market trends for Malaysia's primary agricultural commodities. Combining LSTM, RNN, and ARIMA models to enhance the precision of market forecasts.
3) An online network designed to enhance market connections for farmers, allowing them to reach wider markets and secure more favourable rates through negotiation. This platform offers functionalities for listing, searching, and trading agricultural items, together with integrated real-time price data and comprehensive product characteristics.

2. Background study

The trade of agricultural products has greatly strengthened the global economy. A country generates significant revenue from the commerce of agricultural commodities. Authorities are prioritizing methods to broaden their markets and improve their export endeavors. Countries are simultaneously giving priority to the well-being of their local businesses. Agriculture in Malaysia continues to play a crucial role in driving economic development, holding the third position in terms of importance. Agriculture contributes to economic development by fostering the demand for non-agricultural goods and providing food and raw materials to other industries (Dardak, 2020).

Deep Learning (DL) techniques, which include recurrent neural networks (RNN) and convolutional neural networks (CNN), have been the subject of a substantial amount of research over the course of the past few years. These techniques have also been implemented in a wide range of industries, including agriculture (Maha Altalak et al., 2022). When it comes to the capacity of various DL techniques to gather and assess data, as well as to assist farmers in making the most appropriate decisions at the most appropriate moment, there are variations that may be found. It is possible to discover solutions to the challenges that farmers are facing when performing agricultural tasks by providing them with information, knowledge, and skills. For this purpose, it is possible to address the hurdles that farmers are confronting. According to Nadia (2020), this is due to the fact that it has a strong capacity to address these difficulties.

According to Sun and Zhang (2018), one of the most difficult challenges in the field of time series and machine learning is attempting to forecast the direction that the financial market will take. This is one of the most challenging problems. Sezer et al. (2020) have attempted a variety of approaches over the course of the previous few decades in order to successfully develop an automated financial trade decision-making system and anticipate financial data. These approaches have been attempted in order to get positive results.

It is projected that the implementation of ATP, which employs deep learning and artificial intelligence in the process of data visualization and market trend forecasting, will have a significant impact on the agricultural commerce that takes place in Malaysia. In order to estimate the price market trends of agricultural commodities, approaches that are based on deep learning, such as LSTM, have been utilized. According to Murugesan (2021), these techniques have been demonstrated to be quite effective in terms of producing accurate forecasts and lowering error metrics. In addition, the application of techniques that deal with machine learning can be utilized in order to identify certain agricultural commodities that are being produced and consumed in specific places. Farmers are able to create more effective market ties as a result of this, which in turn helps them to reach a larger range of markets and negotiate higher prices.

2.1. Agriculture Trading Platfrom (ATP) techniques

Deep learning-based models can be classified into three fundamental approaches when applied to ATP systems. These approaches are as follows: Long Short-Term Memory (LSTM) for predicting agricultural commodities prices, Recurrent Neural Networks (RNNs) for analysing temporal patterns in crop yield prediction, and Convolutional Neural Networks (CNNs) for mapping different types of crops. By implementing models that are appropriate and well-suited into the ATP system, the organisation has the ability to improve the user's experience of obtaining data output that is reliable and precise.

2.2. Deep learning

A neural network that consists of three or more layers is the foundation upon which a deep learning framework is constructed. Neural networks and artificial intelligence are two fields that are relevant to this paradigm. Deep learning is a computational approach that makes use of several nonlinear transformations in order to represent complex ideas contained inside data (Dargan, 2020). Deep learning has a number of advantages, one of which is feature learning, which refers to the process of independently extracting characteristics from data that has not been processed.

2.3. Long Short-Term Memory (LSTM)

1) Related work of LSTM

There is a major dearth of research that has employed deep learning strategies for the purpose of agricultural price forecasting. The research conducted by Sabu and Kumar in 2020 reveals that the utilisation of deep learning techniques for predictive analytics makes it possible to conduct an in-depth analysis of even brief historical data pertaining to agricultural prices. To illustrate the architecture and data flow of the LSTM model, **Figure 1** presents a high-level representation of the model's operational structure. Through the implementation of this strategy, it is envisaged that the concerns of all stakeholders would be addressed. According to Elsheikh et al. (2019), it is reasonable to declare that the study's findings justify the use of novel techniques derived from deep learning for the purpose of prediction. It is anticipated that the current agricultural scenario would improve and supply farmers with crucial information regarding the Minimum Support Price (MSP) that is best for their crops given the current circumstances. In addition to this, it has the potential to act as a central hub where both buyers and sellers can investigate a variety of possibilities and make judgements that are appropriate (Rakhra et al., 2021).

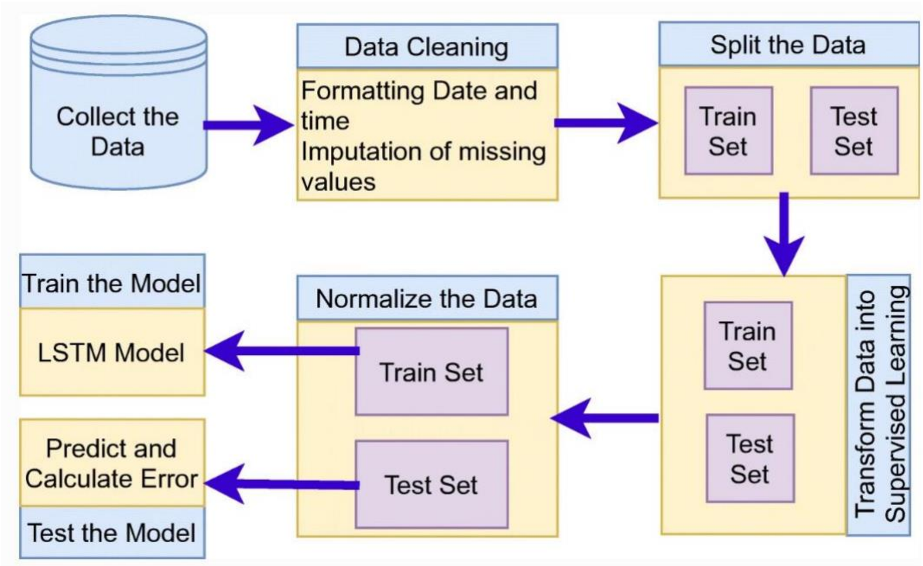

Figure 1. Flow of LSTM model (Banerjee, 2022, p. 15).

There have been numerous suggestions made over the course of time regarding the use of machine learning techniques for the aim of predicting agricultural prices. On the other hand, there is a dearth of studies about the effectiveness of machine learning in predicting the future values of agricultural commodities. Because of the complex and unpredictability of the process, there is a lot of difficulty involved in making predictions about agricultural prices. The reason for this is mostly due to the fact that these costs are strongly impacted by a number of factors, notably those related to the environment.

According to Dariya et al. (2021), there has been a substantial amount of advancement in artificial intelligence (AI) in recent times, notably in the field of deep learning (DL), which is an essential component of AI. The process of deep learning involves the automatic identification and extraction of significant features from the data. When compared to traditional machine learning techniques, deep learning algorithms provide a number of advantages. These algorithms do not rely on the manual production of information. A comparative analysis of deep learning price prediction models was carried out by Nassar et al. (2020) in their research. They compared these models to eight statistical and benchmark machine learning methods. As part of their investigation, the researchers utilised time series datasets that contained information on flowers, fruits, and vegetables. The findings of the study indicated that neural network models, more especially the LSTM algorithm and the

CNN-LSTM, are effective in accurately predicting the prices of fresh fruit with a lead time of up to three weeks.

As of the year 2020, Sabu and Kumar have a paper published. Applied analysis of time series and predictive machine learning approaches were utilised in order to make an accurate prediction regarding the monthly pricing of arecanut in the state of Kerala, which is located in India. During the course of their investigation, they discovered that the LSTM neural network exhibited a higher level of efficacy in this prediction evaluation. Weng et al. (2019) conducted a study with the objective of determining whether or not ARIMA and Deep Learning models are efficient in predicting the values of agricultural commodities by making use of datasets that were collected at daily, weekly, and monthly intervals. The deep learning methodology was shown to be the most reliable approach for forecasting agricultural commodity prices, according to the findings of their study report. A reference to this discovery can be found on page 4.

In an independent investigation that Chen et al. (2021) carried out, they made use of Wavelet Analysis (WA) in order to reduce the amount of noise that was present in the cabbage data. Following that, the data that had been modified and standardised was incorporated into the LSTM model, which ultimately led to improved results in terms of accuracy. The research that was carried out by Namini et al. (2019) reveals that the LSTM model is superior to the This conclusion was arrived at by studying the results obtained. In light of this, it is essential to make use of models such as LSTM, which have the capability of unearthing concealed patterns and dynamics within enormous volumes of data, in order to achieve effective forecasting. The Long Short-Term Memory (LSTM) model proves that it is capable of efficiently managing increasing data volumes and more data complexity in comparison to existing models. Furthermore, it is projected that over the course of a longer period of time, the data that has been collected will continue to expand in both amount and complexity. Yuan (2020) went through a series of tests in order to identify an algorithm that possesses exceptional precision.

2.4. Recurrent Neural Networks (RNN)

1) Related work of RNN

The profitability and livelihoods of farmers are adversely affected by the fluctuation in rice yield from one season to the next, as demonstrated by a Neural Network Model (Shruti Kulkarni, 2020). Farmers and other stakeholders will find it simpler to make critical agronomic and crop selection decisions if they can enhance their capacity to predict crop productivity in a variety of climatic conditions. The utilisation of neural networks enables the prediction of rice production and the examination of the factors that influence the production of rice crops in various districts. Diverse criteria were assessed during the Kharif season, which spans from June to November. The criteria included the process of precipitation, the lowest temperature, the mean temperature, the maximum temperature, the standard crop evapotranspiration, as well as the area, creation, and yield. The dataset was processed using the WEKA tool. The author employed Artificial Neural Networks to evaluate the efficacy of the Artificial Neural network and to generate predictions regarding the yields of rice crops. Ransom et al. (2019) conducted a study to evaluate the effectiveness of machine learning methods in the development of maize nitrate suggestion instruments. For this objective, the investigation implemented sediment and climate data.

Random forest and multiple linear regression were implemented as machine learning methodologies in the investigation conducted by Shahhosseini et al. (2019) to forecast the quantity of maize and nitrogen residue that would be generated. Kim et al. (2019) conducted a study in which they employed a deep neural network model to predict the amount of crop yield between 2006 and 2015. The input parameters for the model were optimised and derived from satellite data and weather datasets. In 2020, Saeed Khaki et al. conducted a study that demonstrated a method that was significantly more effective than popular methods such as DFNN and random forest analysis. The paradigm that has been proposed is a combination of RNNs and CNNs. The CNN component of the model was designed to capture the internal temporal dependencies of meteorological data and the spatial dependencies of soil data collected at varying depths. The RNN component of the model was developed to accurately represent the upward trend in food production that has been evident for several years, as a result of the continuous advancements in plant breeding and management. The success of the model was dependent on a variety of factors, including the season, soil, and possibly the weather and the management of the model. The model that was developed was capable of accurately predicting yields in scenarios that had not been previously tested. This implies that it has the potential to be employed in the future to forecast yields.

2.5. Convolutional Neural Networks (CNN)

1) Related work of CNN

Figure 2 below demonstrates how CNN-based models are applied to map agricultural trends based on data inputs. It is the aspiration of every farmer to optimise their crop yield and derive all possible benefits. One effective method is to evaluate the potential produce of each field in order to strategize your harvest. ‘Sensing and mapping’ is a fascinating development in this discipline. Field yield mapping is the process of generating a map of a field’s productivity by utilising digital image processing and imaging techniques. Precision agriculture is fundamentally dependent on yield mapping.

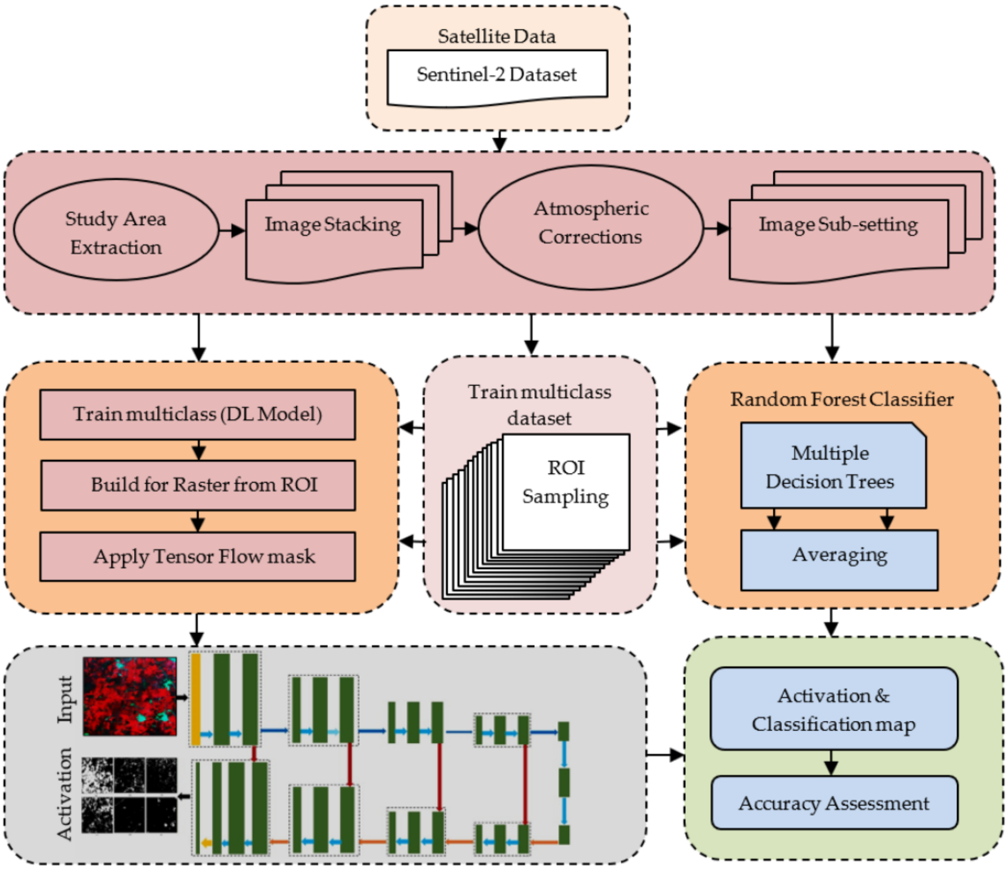

Figure 2. Operational model for the mapping for agriculture land using ENVINet5-deep learning (Singh, 2022, p. 15).

It is designed to emphasise the discrepancies in soil composition that exist among various regions of the farm. Mapping also enables the farmer to address a variety of land-related challenges by providing information on moisture content. The utilisation of machine learning algorithms to analyse historical data is a well-established approach to the generation of yield maps. Farmers can identify the regions that are most and least suitable for crop cultivation by utilising this approach. Additionally, this application has the capacity to generate additional insights by integrating a variety of sensors, including header position sensors and grain flow sensors. Yield mapping is a supplementary instrument that aids in the optimisation of fertiliser utilisation.

Mazzia (2020) developed a novel deep learning model known as Pixel R-CNN by utilising multitemporal decametric Sentinel-2 data from middle northeast Italy. This model was employed to categorise commodities and terrain. The architecture proposed that is based on Pixel R-CNN is significantly superior to other well-known methods, as it obtains an overall precision of 96.5% with a kappa value of 0.914 for 15 subclasses. One of the most exciting aspects of our design is its ability to autonomously extract features by analysing the relationship between various images over time. This capability reduces the necessity for manual feature engineering and enables the modelling of phenological stages in crops. Yang et al. (2019) conducted a study that investigated the utilisation of Convolutional Neural Networks (CNN) to predict the quantity of rice that will be produced by analysing images captured from space. The yield was accurately predicted by the CNN model at every stage of the hardening process, as indicated by the data. Khaki (2019) employed deep CNNs to predict a decrease in maize yield across 1560 locations in the United States and Canada. Deep learning-based change detection has been employed to detect and map agricultural seasonal fluctuations using Sentinel-2 data. Remote sensing is both economical and efficient for monitoring agriculture on a large scale. The use of high-resolution satellite datasets enables the more efficient and effective mapping and monitoring of agriculture. Deep learning is increasingly being employed in a diverse array of scientific disciplines as a result of the widespread availability of powerful computers. Deep learning (U-Net) has been implemented to map the diverse forms of agricultural land usage with Sentinel-2 (Singh, 2022).

2.6. Autoregressive Integrated Moving Average (ARIMA)

1) Related work of ARIMA

The ARIMA (Autoregressive Integrated Moving Average) model is a commonly employed time series forecasting technique in the field of agriculture forecasting as shown in **Figure 3** below. The ARIMA model is utilised to analyse, depict, and investigate the changing patterns of time series, forecast its future trajectory, and establish the matching time series model. The ARIMA model is superior to other forecasting models due to its ability to anticipate without the requirement for additional relevant exogenous variables. It primarily focuses on analysing the patterns and trends of the variables under study. The ARIMA model has been utilised to forecast the prices of different vegetables in relation to the impact of COVID-19 (Mao et al., 2022), agricultural pricing (Jadhav et al., 2018), the price of medium quality rice (Margaretha et al., 2018), and agricultural production (Senthamarai Kannan et al., 2020).

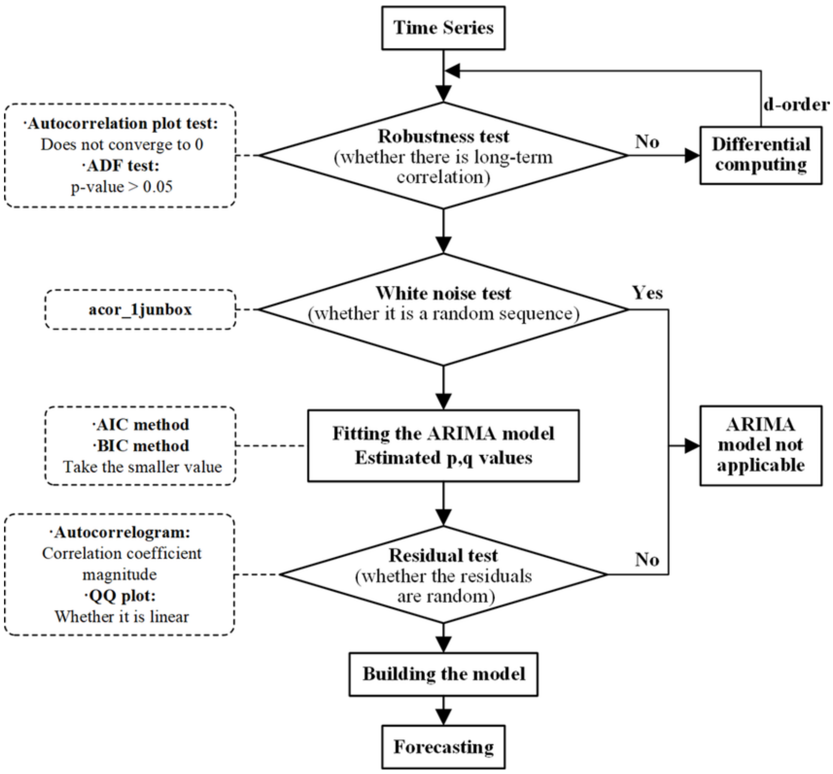

Figure 3. O ARIMA model flowchart (Mao, 2022, p. 8).

Azhar et al. (2020) employed the Autoregressive Integrated Moving Average (ARIMA) theory and an empirical technique to simulate the daily stock values of companies listed in the Jakarta Islamic Index (JII). Chandraprabha and Dhanaraj (2020) conducted a comparative analysis of different time series forecasting techniques to anticipate agricultural production. The study highlighted the effectiveness of ARIMA in precisely capturing temporary fluctuations in yield patterns. Chandraprabha and Dhanaraj highlighted the use of ARIMA for predicting environmental variables, showcasing its effectiveness in capturing recurring patterns. Research using time-series data for short-term forecasting has shown a relatively small margin of error (Virginia et al., 2018). Mao et al. (2022) employed the ARIMA model to predict agricultural prices, with a specific emphasis on the pricing of various vegetables. The ARIMA model was used to examine and clarify the dynamic characteristics of time series, predict future patterns, and construct relevant time series models. The results indicated that the estimated values closely matched the observed trend, and the difference in prices was successfully controlled within a particular range. Joseph and Bahati (2023) investigated the application of the ARIMA model for forecasting maize output in Tanzania. The results suggested that the ARIMA (1, 1, 1) model was the best option for forecasting maize production in Tanzania. The study used time series data of maize yield from 1961 to 2021 and applied the ARIMA technique, known for its reliable forecasting and predictive skills.

2.7. Summary of literature review

For each of the four deep learning-based model techniques for Agriculture Trading Platform (ATP), numerous pertinent works have been assessed. For each technique, an average of 5 related publications have been reviewed. The papers' method, advantages, disadvantages, and improvement have all been covered and explained as shown in Table 1. The related works utilising Long Short-Term Memory (LSTM) Model, Autoregressive Integrated Moving Average (ARIMA), Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are covered in the table as below.

Table 1. Summary of related works.

Article/Journal Paper	ML used	Advantage	Disadvantage
LSTM is a straightforward deep learning method used to estimate market prices for fresh fruit (Azhar et al., 2020).	LSTM	ability to accurately forecast Fresh Products (FP) prices for a period of three weeks in advance.	Machine Learning (ML) and basic Deep Learning (DL) models exhibit lower accuracy in price prediction without incorporating attention mechanisms.
LSTM neural network in forecasting prices (Chimmula & Zhang, 2020).	LSTM	Assist farmers in optimising resource allocation to enhance crop yield and maximise profitability, leveraging price forecasts.	The limited availability of training data poses a challenge for utilising LSTM, a deep learning technique that demands a substantial amount of data.
The Autoregressive Integrated Moving Average (ARIMA) model was employed to predict the price of agricultural items (Mao et al., 2022).	ARIMA	The ARIMA model is suitable for accurately predicting periodic and continuous long-term data.	Demands consistent and regular data on a certain plant kind from a single location.
An automated agriculture system that utilises unique machine	LSTM	Government officials and farmers may benefit from more efficient plantation	The insufficient data size hindered the model's ability to provide accurate

learning approaches to predict commodity prices (Sabu & Kumar, 2020).		techniques if agricultural commodity price data were easily accessible.	predictions that align with the system's requirements.
A comparative analysis of the predictive capabilities of the ARIMA method and models using LSTM for historical data (Siami-Namini et al., 2018).	ARIMA LSTM	Promotes the advantages of utilising deep learning-based algorithms and techniques for analysing economic and financial data.	There are other additional forecasting challenges in the fields of finance and economics that can be addressed through the application of deep learning techniques.
Agricultural commodity price prediction application based on the Long Short-Term Memory (LSTM) model (Sabu & Kumar, 2020).	LSTM	Taking many economic and environmental factors into account makes it easier to guess what prices will be in the future, giving forecasters more freedom.	Insufficient comprehensive and localised price analysis is hindering the ability to facilitate trade between farmers involving agricultural items.
An assessment was conducted to examine statistical and machine learning techniques for integrating soil and weather data into nitrogen recommendations for maize (Naorem et al., 2019).	Random Forest	The random forest models demonstrated superior ability in enhancing all three N recommendation systems.	Regrettably, not all Machine Learning (ML) techniques have been successful in enhancing all N recommendation programmes.
Evaluated four machine learning techniques as meta-models for predicting simulated maize yield and N loss (Guevara-Viejó et al., 2021).	Random Forest	Assist in expediting the advancement of integrating simulation models and machine learning (ML) in order to provide dynamic decision support tools for pre-season management.	The use is restricted due to the data prerequisites and extended execution times.
An analysis of prominent artificial intelligence models for predicting crop yields (Saprabha & Dhanaraj, 2020).	DNN	The Deep Neural Network (DNN) utilising the July-August JA model has the capability to properly predict maize and soybean yields for a specific year far in advance.	Lack of further data for training.
Categorization of crop resilience to high temperatures and water scarcity (Manogna, 2020).	Support CNN	exhibit greater yield stability in comparison to non-tolerant hybrids under these stress conditions.	Under extremely harsh and demanding circumstances, all crops would ultimately wither and deteriorate.
Pixel Recurrent Convolutional Neural Network (R-CNN) (Virginia et al., 2018).	R-CNN	capability of automated feature extraction by learning time correlation of multiple images, which reduces manual feature engineering and modelling crops phenological stages	direct quantitative comparison of the classification performed in these studies is difficult due to various dependencies such as the number of evaluated ground truth samples, the extent of the considered study area, and the number of classes to be evaluated
Utilising unmanned aerial vehicle (UAV) captured remote sensing photos, this study employs deep convolutional neural networks to accurately estimate rice grain yield throughout the ripening stage (Weng et al., 2019).	RNN	Data augmentation methods, including adaptable cropping at random and rotation, helped with the overfitting issue.	Various management approaches can induce stress on biological components, which in turn can impact the progression of phenological stages.
Challenges and emerging opportunities in the agricultural sector of India (Naorem et al., 2019).	KNN	Enhancing the nutrient usage efficiency of fertilisers, mitigating soil deterioration, and improving soil quality are achieved by complementing the current technologies, resulting in a reduction in environmental pollution concerns.	Limited internet access in rural areas, expensive service fees, and lack of basic computer literacy and education hinder the rapid development of e-agribusiness.

2.8. Existing systems

The systems that have been assessed are AgFlow, IndexBox, and Agri-Food Data Portal. These systems that are currently accessible are pertinent to the specific field, and a thorough analysis has been conducted on the many aspects and characteristics of these systems. Included among these features are insights into market prices for agricultural commodities, a prediction function for market trends, and information on current production and consumption trends as well as listings.

1) AgFlow

AgFlow at https://www.agflow.com is a platform that provides agricultural market data and analysis to help traders and businesses make informed decisions. It offers a range of services, including market analysis, physical market data, and agricultural commodity trading tools. The platform collects raw data from a network of sources, aggregates it through proprietary technology, and redistributes it across the entire agricultural value chain.

AgFlow offers users real-time access to cash price quotes for more than 75 agricultural commodities. Additionally, it allows users to monitor worldwide demand and supply projections and provides market insights and trading signals. The platform additionally provides an Application Programming Interface (API) that assists customers in constructing dependable predictive models. AgFlow functions as a vital tool for traders and enterprises seeking to stay updated on the most recent advancements in the agricultural sector and make decisions based on data.

AgFlow offers users real-time access to cash price quotes for more than 75 agricultural commodities. Additionally, it allows users to monitor worldwide demand and supply projections and provides market insights and trading signals. The platform additionally provides an Application Programming Interface (API) that assists customers in constructing dependable predictive models. AgFlow functions as a vital tool for traders and enterprises seeking to stay updated on the most recent advancements in the agricultural sector and make decisions based on data.

In summary, AgFlow is a platform that provides agricultural market data and analysis to help traders and businesses make informed decisions. It offers a range of services, including market analysis, physical market data, and agricultural commodity trading tools. While the platform has several strengths, such as comprehensive coverage and real-time market data, there are also potential areas for improvement, including user interface and subscription cost.

Weaknesses:

- Need for a more user-friendly interface.
- High cost of subscription.
- Limited coverage of certain niche agricultural commodities.

2) IndexBox

IndexBox at https://app.indexbox.io/ is a market intelligence platform that provides businesses with data-driven insights to help them thrive in today’s global marketplace. The platform offers access to market data on over 10,000 products across 200+ regions, with the aim of empowering companies to make informed decisions and identify new opportunities. IndexBox specializes in providing businesses with data analytics, market research, and insights through its AI-driven platform. **Figure 4** shows the IndexBox production trends and exports.

Figure 4. Snapshot of IndexBox production trends and exports.

Index Box's primary features include the capacity to track data sources and prediction models, digital access to data on a variety of products, and machine learning models that can estimate market growth, demand, and prices. IndexBox also provides online access to data on a variety of products. Aside from that, the portal provides access to online training sessions and gives users the choice to download files for free or for credits. In order to guarantee a high level of data integrity, the algorithms that make up IndexBox make use of a variety of information sources for cross-checking and validation. IndexBox is a powerful and easy-to-use market intelligence platform that provides businesses with data-driven insights to help them succeed in the global marketplace. The platform offers a wide range of features, including AI-driven data analysis, forecasting, and online training sessions, making it a valuable resource for companies looking to make informed decisions and identify new opportunities. While the platform has several strengths, such as comprehensive coverage and AI-driven data analysis, there are also potential areas for improvement, including user interface and pricing.

Weaknesses:

- Need for a more user-friendly interface.
- Cost of accessing certain features or reports

3) Agri-Food Data Portal

The Agri-Food Data Portal of the European Commission provides information on agricultural markets throughout Europe and the world. Among the many resources and tools offered by the platform are visual representations of data about output, prices, imports, exports, and aid programmes. An API is available on the site, allowing users to retrieve data from the database, and a machine-to-machine interface has been added to the agricultural markets portion. The platform's stated goal is to facilitate evidence-based decision-making in the agricultural industry by providing thorough and trustworthy data.

The main features of the Agri-food Markets section include market data on various agricultural sectors, such as beef, meat, poultry, eggs, goat, wine, dairy, vegetables, cereals, fruits, rice, oilseeds, protein crops, sugar, olive oil, and sheep. There are detailed views of various animal types and quality of meat classes on the platform, as well as data on live piglets and various pork types and grades. Additionally, the platform provides detailed monthly trade data, weekly PDF dashboards, and dedicated apps for prices of milk and dairy products. The impact of the platform lies in its ability to provide policymakers, researchers, and stakeholders with access to high-quality, standardized data, ultimately supporting informed decision-making and the development of sustainable agricultural policies.

Weaknesses:

- Provides access to market data on various agricultural sectors
- Data integration and standardization across different agricultural domains may present challenges
- The platform may cost from additional features to enhance data visualization and analysis

2.9. Summary of existing systems

The aforementioned functionalities are not available in all systems. Some examples are AgFlow, IndexBox, and the Agri-Food Data Portal. The suggested system will have all the elements that are needed, such as a vendor directory, real-time market price insights, a commodity info page, a commodity forecast page, and weather information. For the suggested system to address the problems mentioned earlier, several elements are necessary.

3. Requirement analysis

3.1. System overview

There will be two types of users in the system, which are user and admin. The user and admin land on the homepage, where they are presented with an overview of the platform’s features and options. Next, users who are new to the platform need to click on the “Sign Up” button and complete the registration form. Admin will verify the user registration application. After that, the system will send a verification email. Next, after receiving an email requesting confirmation of their email address, users must click the link to finalise the process. After verification, they are redirected to the platform. Returning users click on the “Log in” button where they insert their email and password. Next, users are taken to their personalized dashboard, which provide overview of recent activity and manage personal information. After that, users are prompted to complete their profiles. users provide details about specify their products interests, preferred categories, and location to be search.

User also can click on agriculture commodities info page where it will show the current real-time agriculture information like product details, product price, weather information, vendors directory based on what user insert in the search filter bar such as commodity type and location. Once a use selects or choose the specific location like state and districts, the user may view the list of various agriculture products that available on that area and once the user selects the product it will show the list of registered and verified individuals or vendors that produce, distribute, or supply the product. Next, user can easily choose their preferred vendors and it will direct the user to the vendor individual information page where it will contain the vendor's details like their farms, products, farming experience and contact information. Admin can edit previous data in database by adding or removing the data and able to update new data to database.

Apart from that, user can click on commodity forecast page option and the screen will show list of location in Malaysia to be choose. Next, once the user chooses specific state and district, then user need to specify the duration prediction then the screen will appear list of agricultural products available within the timestamp preferred either monthly or annually. Next, user will need to choose specific commodity type to be forecasted. After that, the system will preprocess data for forecast process than visualize the forecast result. Therefore, user will be able to view forecast result in graph for production and consumption. The system will store and retrieve data from the database and user will be able to view back their recent activity history. In addition, user receive platform updates and able to access help centre with FAQs, guidance, and customer support team options.

3.2. System requirements of ATP

Functional requirements:

- The system must provide the functionality for users to register and authenticate their login credentials.
- The system is required to predict the future pricing of agricultural goods.
- The system should visualize the forecast result clear for users.
- Users shall be able to select interested commodity.
- System should store and retrieve data in a database.
- Commodities data shall be downloadable.
- A user profile shall be created automatically upon registration.
- System shall be able to authenticate existing user diagram.

Non-functional requirements:

- The system shall be able to access the system anywhere with an internet connection The system shall be able to handle user queries do not exceed 30 s when backend is running.
- The system shall be simple to access for a first-time user. The redirecting page amount for any feature of the system shall not exceed 5 pages.
- The forecasting result shall be available for at least 5 commodities. For each commodity, univariate model, and multivariate model for forecast process shall be able to be accessed. The meantime of download a file in csv format from a software shall not exceed 10 s.
- The system would require user account to access to all features. User shall register their account with their email address one time only, system will reject account registration of existing account
- Commodities listed in the software shall be in English. All commodities prices in the software shall be in Ringgit Malaysia, and local pricing for every commodity

- A web-based interface allows users to access the ATP through a web browser. It provides a user-friendly interface that can be utilized from various devices, including desktops, mobile phones, iPad, smartphones, laptops, and tabs. The web-based interface should be responsive and adaptable to different screen sizes and resolutions

3.3. System specifications

3.3.1. Personal laptop

- A CPU with a minimum clock speed of 1 GHz, either 32-bit (x86) or 64-bit (x64).
- Requires a minimum of 2 GB for 64-bit systems and 1 GB RAM for 32-bit systems Windows 10 or equivalent or MacOS.
- Intel Core i5 or equivalent.
- 256 GB SSD or higher.
- Compatible internet browser (e.g., Chrome, Safari, Edge).
- Internet connection for accessing the web-based platform.
- Mouse and keyboard for navigation.
- Screen for viewing the web-based interface.

3.3.2. Server

- The recommended operating systems for the application server are either Windows Server 2019 or Windows Server 2022.
- The processor options for this device are the 11th Gen Intel Core i5-11400H running at a base frequency of 2.70 GHz, the Intel Xeon E5 or E7 series, or the AMD EPYC series.
- 500 GB SSD for the operating system and applications, additional storage for data based on the platform’s expected scale.
- High-speed internet connection with at least 100 Mbps upload and download speeds.
- The version of MySQL being used is 8.0.31 for Win64 on x86_64. It is the MySQL Community Server under the GPL license.

3.4. Use case diagram

For a first-time user, it is necessary for them to create an account by clicking Sign up option. After completing the registration, they can start using the system by logging in their details. User can manage their personal details in the dashboard where they able to edit and update it. There is an option to view agriculture commodities forecast where the user needs to choose the forecast location (specify state or district), then user choose the forecast duration (next week/month/year), then choose the commodity type available based on the location inserted. Then once user complete all the forecast requirements, the system will visualize the forecast graph result. Next, user also able to view commodity information like current price, production, consumption and more. In addition, user can view weather information details and vendors directory which contain vendors details like their farms, products, farming experience and contact information. Thus, user able to reach FAQs and support guidelines page where user able to view needed information. Lastly, user able to logout once done use the platform.

Administrators no need to register for an account since it would be provided to them. They just need to login in order to access the platform. On the admin interface, they able to verify user registration whether get approved or rejected based on requirements. Admin can edit previous data in database by adding or removing the data and able to update new data to database.

The system is an external system which will preprocess data for forecast process which required by the user. Then system will visualize the forecast result in graph form for easier reading and analysing purposes. The system also will train the forecast model on doing predictions for agriculture commodity market price stock, production and consumptions. Lastly, the system will store and retrieve data from database.

3.5. Context diagram

The context diagram, depicted in **Figure 5**, aids in discerning the transmission of information between the system and its external environment. The diagram also illustrates the limits of the system. The system functions as a trading platform specifically designed for agricultural use. The exterior setting consists of the system's users, specifically user and admin. The data is transmitted from the source to a specific target location, and this process is visually represented in the diagram using arrows.

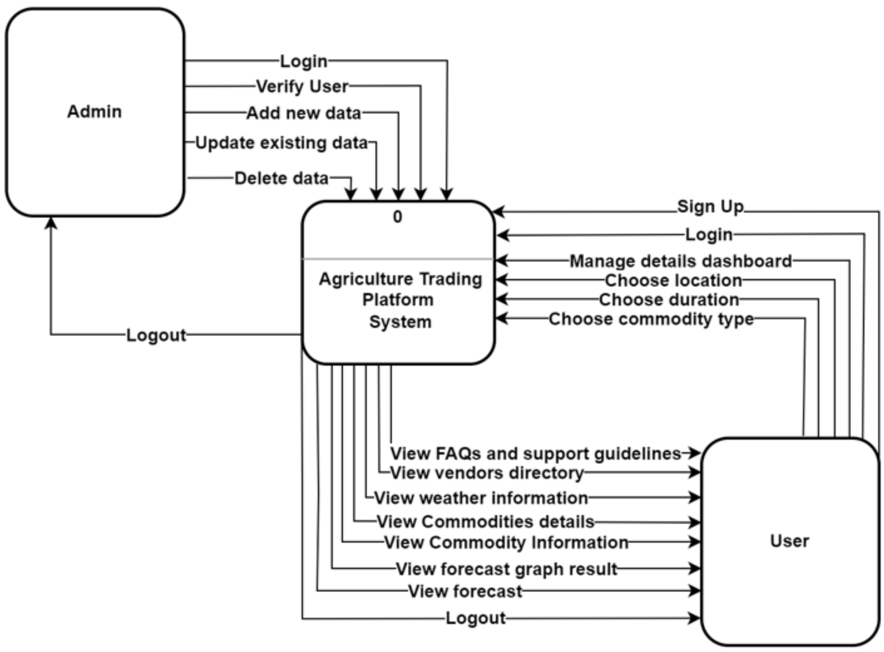

Figure 5. Context diagram of the ATP.

4. Design

4.1. System architecture

In this chapter, the software architecture, sequence diagrams, screen designs and the data dictionaries are illustrated and described. **Figure 6** illustrates the system's architectural architecture of the program. The software's forecasting feature utilizes two distinct types of data in its datasets: past data and current data. The forecasting model is trained using past data collected during the construction process. The implementation of the software's forecasting models is facilitated by real-time data, which offers regularly updated information. While the project is being built, the models are taught. If current data makes up 2/3 of the past data, the models are re-trained in the post-development phase. Historical data refers to real-time data that is utilized for model training. These assurances ensure that the precision and efficiency of the model will be regularly improved.

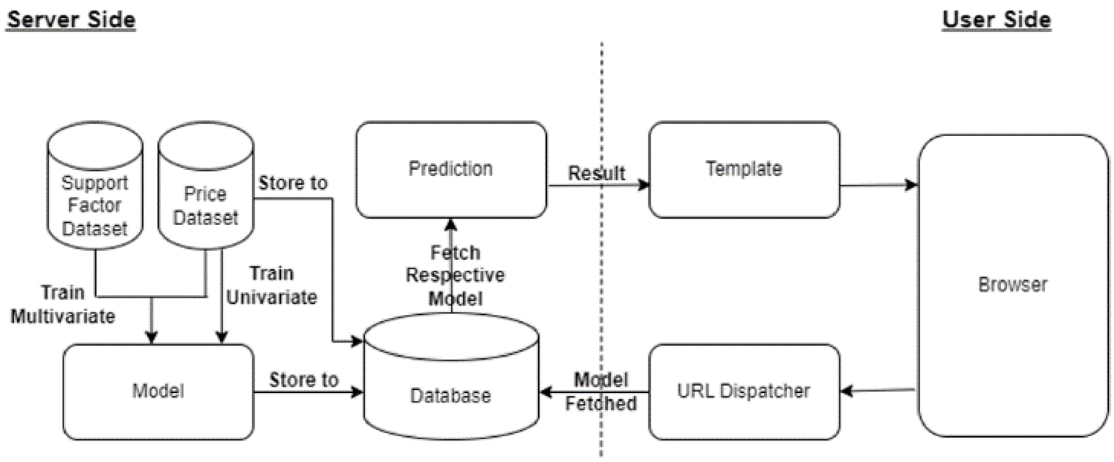

Figure 6. LSTM forecast model system architecture.

The software's datasets are classified into two categories: supporting factor datasets and price datasets. They are employed in the training of both univariate and multivariate models. Training multivariate models involves utilising a combination of supporting factors and price data, whereas univariate models solely rely on the price dataset. Specific supporting variables are selected for each commodity based on their link with the commodity prices. The system examines the comparison between both models of the same commodity, and occasionally, the accuracy of the univariate model surpasses that of the multivariate model. During the development phase, both univariate and multivariate models are trained and saved in the database for implementation. During the prediction process, the stored trained model will be retrieved from the database. The prediction result, along with the historical data, will then be displayed in both graph and tabular formats results.

4.2. Data flow architecture

4.2.1. Data source collection

Pricing dataset has been collected from Federal Agricultural Marketing Authority Malaysia (FAMA) and Malaysia Department of Agriculture. The collected data provides a month-wise commodities market insights and price report for each vegetable for every state and district. The monthly data for short-term forecasting and annually data for long-term forecasting. The registered farmers and vendors list dataset is collected from the website Jabatan Perkhidmatan Veterinar "https://www.dvs.gov.my/dvs/resources/user_1/2019/BDKK/12.11_.2019-SENARAI_KESELURUHAN_mygap_LADANG_TERNAKAN_.pdf". The production of fruits and vegetables dataset is collected from the website official portal Malaysia Department of Agriculture "https://www.doa.gov.my/index.php/pages/view/622?mid=239". The market price, production and consumption area dataset obtained from Federal Agricultural Marketing Authority Malaysia (FAMA).

4.2.2. Data preprocessing

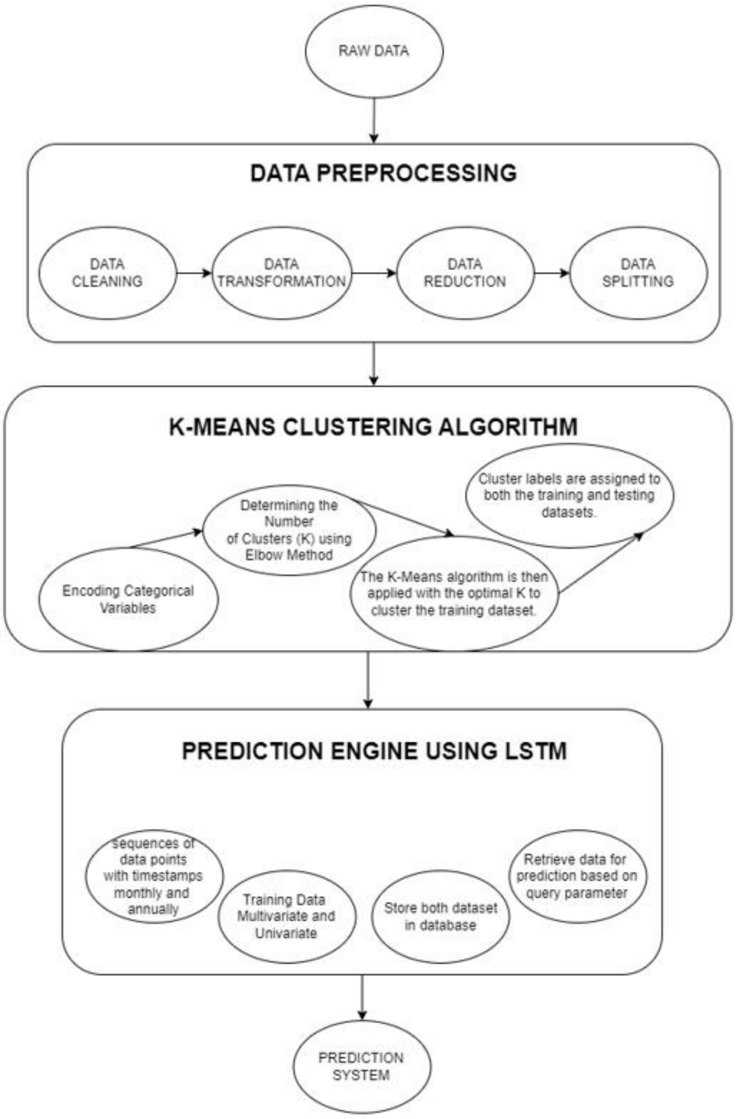

Figure 7. Dataset system architecture.

Figure 7 illustrates the flow for system architecture. The raw data must be filled out before it can be deemed complete, preprocessing is necessary. The term "data preprocessing" refers to the steps taken to make raw data more understandable and useful. There are usually four steps to preprocess data. These include data cleansing, data transformation, data reduction, and data splitting. Normalising a dataset requires data cleansing, which entails removing missing or inaccurate information. The term "transformation" describes the process of mapping the homogeneity of the data. Hourly, monthly, or yearly access to the data is one possibility. The data must be organised by month and year. Data reduction refers to the process of organising data into a simpler format. Lastly, data partitioning: The algorithm partitioned the dataset into training and testing sets for the LSTM and clustering models. The data are structured in an unsupervised manner here. In order to convert the data into a supervised format, clustering is necessary. One of the simplest clustering algorithms is K-means.

4.2.3. Data cleaning

Data cleaning is the systematic procedure of identifying and correcting inaccuracies, inconsistencies, or missing values in a dataset to ensure its reliability and trustworthiness. Data cleaning ensures the correctness and dependability of the dataset for analysis. Data cleaning enhances the quality and reliability of analytical findings by dealing with missing values and correcting errors. It helps maintain consistency and standards across the dataset, making it easier to understand and analyse. Data cleaning is the process of identifying and correcting flaws or inconsistencies in a dataset. As an illustration, the ‘clean_unitprice’ function substitutes blank characters in the ‘UnitPrice’ column with zeros and changes the column to the float data type. This ensures consistency in data type and handling of missing values. The ‘clean_data’ function is utilised to sanitise the complete dataset by implementing the ‘clean_unitprice’ function and other essential data cleansing procedures.

4.2.4. Data transformation

Data transformation is the process of converting raw data into a suitable format for analysis or modelling using mathematical or statistical methods. Data transformation is a crucial aspect of data analysis as it encompasses the process of standardising formats, handling categorical variables, and scaling features. It enables analysts to extract valuable insights from the data and improves the effectiveness of machine learning algorithms. Resampling is a transformation technique that helps to combine data at different frequencies and handle irregular time series data.

4.2.5. Data reduction

Data reduction techniques aim to reduce the dimensionality or size of a dataset while preserving its fundamental characteristics. The algorithm uses quantile-based discretization to decrease the number of categories in the ‘Category’ column, hence reducing the data. This process helps to optimise the dataset and improve the efficiency of subsequent analysis or modelling tasks. Reduction methodologies facilitate the simplification of complex datasets, improving computational efficiency, and reducing storage requirements. By reducing the dimensionality or categories, analysts can focus on the most relevant characteristics and trends in the data, leading to more effective analysis and modelling.

4.2.6. Data splitting

Data splitting refers to the act of dividing a dataset into separate testing and training sets. This division is done to evaluate the performance of machine learning models. The approach uses the ‘train_test_split’ function from scikit-learn to partition the data into training and testing sets, with a specified test size. This allows for the assessment of the model’s performance on novel data and helps prevent overfitting. Data partitioning facilitates the evaluation of the models’ ability to generalise and the detection of overfitting. Analysts can validate the model's performance and ensure accurate predictions on new data by employing separate training and testing sets. Data splitting is crucial for assessing the reliability and robustness of machine learning models.

4.2.7. K-means clustering algorithm

The K-Means algorithm is used to group the data on agricultural goods into groups. The method K-Means is used for grouping without being watched. Agricultural product data is put into groups using clusters. The number of groups is shown by the variable "k". The centroid is thought to be shown by the first two numbers in the first phase. Next, use the Euclidean method to find the distance between each data point and the centre of the group, which is also called the centroid.

$$\sqrt{Distance} =. \sqrt{(x2 - x1)2 + (y2 - y1)2}$$

Determine the cluster centre with the least distance from the cluster among all computed centroids and assign the data point to it. Compute the updated cluster centroid iteratively until it remains constant from the previous iteration. The data is categorised into agricultural commodities, including poultry, vegetables, fruits, and nuts, as determined by the k-means algorithm. The advantages of K-Means clustering include its high processing speed, ability to manage large data sets, low computing cost, and flexibility in adjusting the number of clusters. K-means clustering was implemented to extract clusters from the dataset, which has undertaken feature selection optimisation (Guevara-Viejó et al., 2021). The dataset is initially organised chronologically on a monthly or annual basis. Subsequently, it undergoes a training data procedure that entails the division of the dataset into training and test datasets. The dataset will be subsequently stored in a database. The data will be retrieved from a prediction model that is based on historical data and utilises LSTM when the consumer requests agricultural commodity forecast market trends. K-Means clustering is employed to identify patterns, groups, or clusters in data without the necessity of labelled samples. It is suitable for managing large datasets due to its computational efficiency and scalability. K-Means clustering is frequently employed for the purpose of consumer segmentation, exploratory data analysis, and anomaly identification.

The formula for determining the distortion is:

$$\text{distortion} = \text{sum}\ (||x - \text{mu}\_k||^2)$$

$$\text{for all data points } x \text{ in cluster } k.$$

The distortion is calculated as the total of the squared Euclidean distances between each data point x and the centroid mu_k of cluster k. The centroid of cluster k is represented by mu_k, x is a data point in cluster k, and ||.|| indicates the Euclidean distance.

4.2.8. Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) structure that is capable of acquiring knowledge about long-term relationships in sequential input. The code utilises the 'build_lstm_model' function to create and train an LSTM model. The model is trained using time series features extracted from the dataset, such as 'Category' and 'State'. Once the LSTM model has been trained, it is employed to forecast future values of the ‘Category’ variable based on historical data. This allows for the forecasting of trends or patterns in agricultural transaction data throughout time. LSTM models excel at capturing extended relationships in sequential data, making them very suitable for predicting future values in time series data. They possess the capability to identify complex patterns and variations in the data, enabling accurate predictions of future values. LSTM models are frequently employed in several applications such as natural language processing, audio identification, and financial forecasting.

5. Implementation

Description of solution

Hardware this chapter details the resolution of an Agriculture Trading Platform that can be accessed through a web-based interface. URL for local access: http://localhost:8501. The network URL is http://10.112.68.214:8501. The Agriculture Trading Platform (ATP) is an advanced web application designed to revolutionise the trading of agricultural goods in Malaysia. It addresses the shortcomings of the current approach, namely its inefficiency and lack of clarity. This adjustment is highly necessary. The purpose of the platform is to facilitate the achievement of the objective of developing a dashboard that can accurately forecast the quantity of certain agricultural commodities that will be produced and consumed. The system employs sophisticated deep learning algorithms to forecast market trends and enhance farmers’ ability to establish more effective connections with buyers and sellers.

The ATP features a user-friendly interface that is accessible to a wide range of individuals, including farmers, producers, distributors, and buyers. The system incorporates numerous features and functionalities that enhance the ease and transparency of selling agricultural goods. Users can authenticate their identity by inputting their login credentials on the password page. Novice users can effortlessly create an account on the website by accessing the signup page. The primary homepage provides users with comprehensive information regarding the platform’s capabilities and options. The document contains comprehensive information regarding the commodities, usage predictions, frequency distributions, a vendor list, and data management tools for administrators. Each of these sites is equipped with its own array of tools to assist users in their job.

The forecast consumption area page provides users with analytics and detailed information regarding the utilisation of agricultural items in various regions. The forecast commodities page provides customers with a convenient way to view projected prices for various goods. These guesses are made using sophisticated deep learning algorithms. The page displaying frequency distributions provides a broad understanding of the dispersion of product rates and unit pricing. The vendor directory page displays a variety of providers and their respective locations. The admin data management page allows platform managers to efficiently handle and organise platform-related data.

1) Login page

Three buttons—”Login”, “Forgot Password?”, and “Sign up”—are on the page. The email and password fields are at the top. When the “Login” button is clicked, the code checks the format of the input and the database passwords. If it's right, it shows a word of success and changes the page's state to "Home." It shows an error message if the data format or passwords don't match. If you choose the other two, the page will just say "SignUp" or "Forget Password."

2) Signup page

The webpage provides users with the option to register by entering their complete name, email address, password, state of residence, and phone number. The code uses regular expressions to validate the email and password entered by the user when clicking the "Sign Up" button, and then compares them to the confirmed password. If all validations are successful, the code utilises the save_data_to_mysql function to store user data in MySQL. If the data is successfully stored, a success message and the text "LoginPage" are displayed. However, the redirect functionality is currently commented out. An error notification is displayed in the event of failed validations or the inability to store the data in the database.

3) Homepage

The page where data from a MySQL database undergoes cleaning and pre-processing by the programme, which involves transforming the pricing date into a datetime object, converting the unit price into a float, and calculating the total price. Users have the ability to navigate through the data by utilising a sidebar equipped with filters for state, district, and category. Additionally, they may access a view of agricultural commodities that allows them to choose certain features. Furthermore, a progress bar is available to display the percentage of a target number that has been achieved. In addition, the software has an LSTM model for forecasting product pricing and a distribution histogram that visually represents the frequency distribution of features. Users can also analyse distribution patterns by dividing features into quartiles within the app. The application features a prominent header displaying its name, a sidebar containing filters and options, and a main area that presents data and functions, resulting in a user-friendly interface. The application includes a logo and a style sheet that may be used to personalise its visual design.

4) Forecast consumption area page

The page forecasting tool utilises the specified date to compute agricultural commodity inventories and unit prices. Prior to allowing the user to select a date, the code verifies if there are any NaN values in the Quantity column. The total price and average amount of each product are derived by grouping the data. This formula determines the necessary stock by dividing the total price by the average quantity and the anticipated unit price.

This tool assists users in forecasting future consumption trends by displaying product demand forecasts. It aids in the process of planning and decision-making by presenting information on stock requirements and forecasting commodities prices. This aids supply chain managers in ensuring sufficient inventory levels to meet future demand and anticipate financial requirements. The dashboard is particularly beneficial for agricultural stakeholders because it presents information visually and allows for filtering by date and state as shown in **Figure 8**.

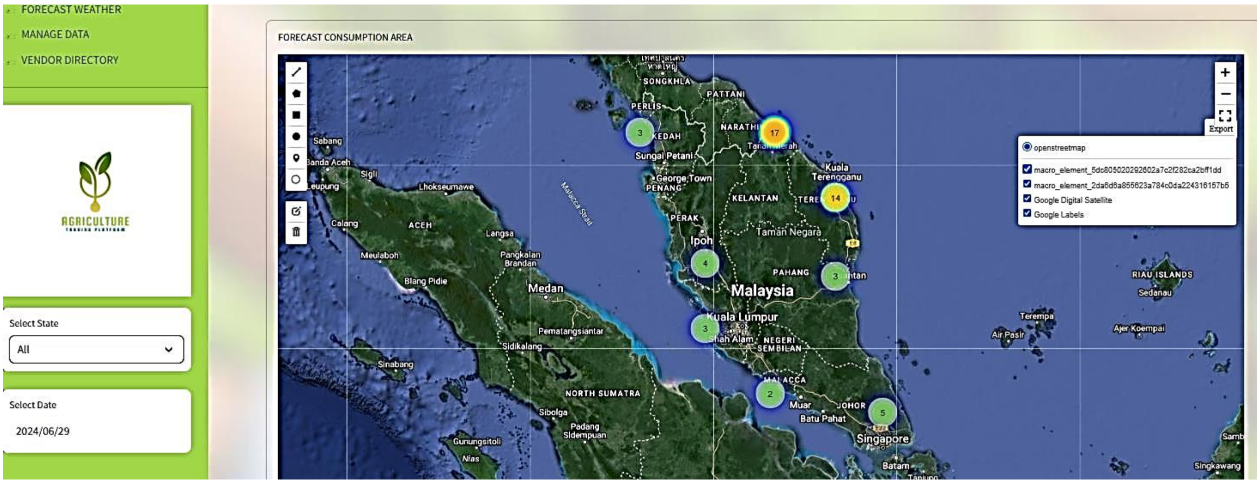

Figure 8. ATP forecast consumption area page.

5) Forecast commodities page

This page utilises historical data to provide a complete forecast of agricultural commodity prices as shown in **Figure 9**. Begin by establishing a connection to a MySQL database in order to fetch information pertaining to categories of commodities, individual goods, states, pricing, and quantities. Streamlit provides customers with the ability to select a category of commodities, a specific product, a state, and a date for forecasting. The data is processed and displayed, ensuring a minimum of 20 data points for the forecast. If the filtered data meets the criteria, the LSTM model is trained using the most recent 20 data points.

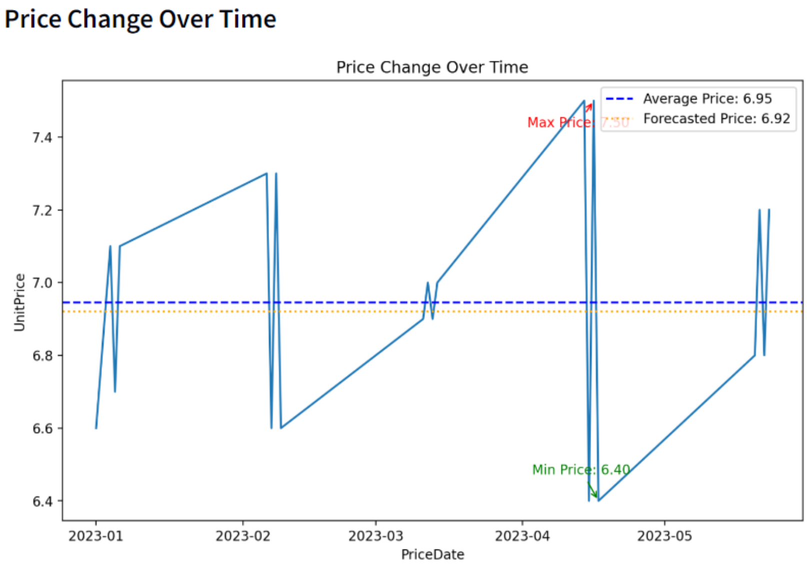

Figure 9. ATP forecast commodity page.

In order to forecast prices, the LSTM model is trained using historical data. MinMaxScaler is a technique used to rescale data in order to improve the performance of a model. The model utilises 10-point sequences to make predictions. The original pricing scale is obtained by performing an inverse transformation on the findings. The projected mean unit price is displayed following the prediction. The code further presents a line plot illustrating the fluctuations in prices over time, encompassing key statistical measures such as minimum, maximum, median, standard deviation, variance, and average prices. The map displays trends and projected data, accompanied by annotations indicating the minimum, maximum, and projected prices.

6) Frequency distributions page

The page generates and displays the unit price and product frequency distributions for agricultural commodities on this page as shown in **Figure 10**. Once the MySQL data is loaded into a DataFrame, the training dataset is shown in an enlarged section of the Streamlit interface. The main features are the calculation of frequency, percentage frequency, cumulative frequency, relative frequency, and cumulative relative frequency for both unit pricing and product frequencies. Users can personalise the condensed information in interactive tables by utilising multi-select choices. The page features a comprehensive graph displaying the distribution of unit prices. Plotly, a robust graphing package created this histogram with a dashed line representing the mean unit price, customisable bar markers, and a well-structured layout including gridlines and a horizontal legend. The histogram displays the distribution of unit prices throughout the dataset, facilitating the identification of patterns and trends. Agricultural commodity frequency distribution study is facilitated by comprehensive data tables and interactive visualisations.

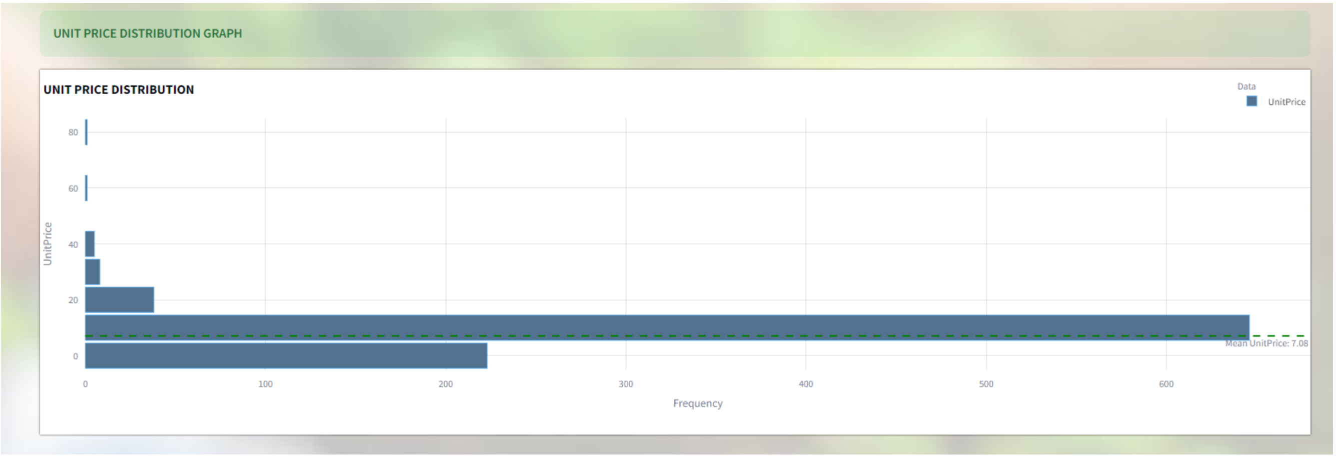

Figure 10. ATP frequency distribution page.

7) Vendor directory page

The website displays and analyses data related to vendor directories and manufacturing areas. The act of choosing suppliers from the sidebar allows for the filtering of data and the presentation of relevant information on a map and graphs. Metric cards display the number of vendors and sales, along with a branding sidebar. The application’s interactive Folium map displays the locations of vendors using clustered markers to enhance visualisation. Every marker triggers a pop-up window displaying vendor details such as management, farm area, productivity, unit pricing, total sales, and contact information. The features of the application include a control for expanding the screen, an overlay that displays a heatmap of vendor density, and tools for annotating the map by sketching it. The user can switch between different map layers, such as Google Maps satellite and label tiles. The application additionally showcases the sales success of vendors using bar and pie charts from Plotly.

8) Admin manage data page

The page displays the administrative dashboard for managing agricultural data. The dashboard collects agricultural data from a MySQL database and allows managers to apply date range filters. The data that has been processed using a filter is presented in an extended section along with metrics that provide a summary of the dataset. Administrators can efficiently analyse the data landscape using these measures, which encompass the overall quantity of agricultural goods, product prices, and total prices. Administrators have the option to utilise dashboard visualisation tools for the purpose of analysing data. Dot plots and bar charts display the costs of products, while scatter plots illustrate the frequencies of product features. Additionally, bar charts depict the quantities of products. The design is minimalistic, including dedicated areas for inputting fresh database entries and side panels for choosing certain time intervals and showcasing a circular emblem. This comprehensive collection of tools assists administrators in the management and analysis of data.

9) Forecast weather

Built with Streamlit, this page is a real-time weather forecasting application. It gets weather information for several Malaysian districts by using the Weatherbit API. The programme is made to deliver users precise and current weather data for a chosen date so they may make decisions based on the weather. A sidebar for choosing the prediction date and a main display area for seeing the weather forecast comprise the simple and user-friendly design. The application opens with a dictionary of coordinates for different districts in every Malaysian state. The software requests the weather data for each district from the Weatherbit API when the user chooses a date from the sidebar. Relevant weather information including temperature, feels-like temperature, pressure, humidity, dew point, cloud cover, visibility, wind speed, and weather description are extracted by parsing the API answers. Following compilation of this information into a list of forecasts, the list is transformed into a pandas DataFrame for display. The software offers weather data in a well-organized tabular style if it is available for the chosen day, so users may easily understand the expected conditions throughout several districts. Should no data be available, the user will receive an error message stating that the chosen date is not accessible for weather data. People and companies who need accurate weather forecasts to schedule their activities will find this software especially helpful.

6. Conclusion

In summary, the existing system for the trading of agricultural products in Malaysia is not very efficient and does not provide as much clarity. Our Agriculture Trading Platform (ATP) is a comprehensive web-based application that was designed with the intention of resolving these issues. By providing a central location for farmers, producers, distributors, and purchasers to list, search for, and exchange agricultural items, the purpose of this platform is to facilitate connections between these groups of individuals. A number of features, including real-time price data, comprehensive product specifications, and improved communication capabilities, are incorporated into the ATP in order to make trading more transparent and effective.

After considering the issue statements and objectives, the ATP was able to effectively construct a dashboard that is simple to operate and provides forecasts regarding the pricing of main agricultural items, the manner in which they would be consumed, and the general movement of the market. For the purpose of identifying and forecasting market trends, the application makes use of deep learning algorithms, which enables users to make more informed decisions. In addition to this, it assists farmers in gaining better access to stores, which enables them to sell their products in larger stores and obtain higher prices.

The ability of the ATP to provide you with real-time analytics, comprehensive forecasts, and complete directories of vendors, all within an interface that is simple to use, is what sets it apart from other similar products. Due to the fact that it makes markets more transparent, minimises information gaps, and makes trading in agriculture more efficient overall, the system has a significant impact on the agricultural sector in Malaysia. Through the process of bringing together various market participants and providing them with essential market information, the ATP assists individuals in making more informed decisions and contributes to the improvement of the economics of the agricultural sector.

Throughout the course of this project, I have acquired a deep understanding that the development of a digital Agriculture Trading Platform (ATP) necessitates a substantial amount of technical and domain-specific expertise, particularly in the areas of Python programming, Streamlit for web applications, and effective data administration. At the outset, it was imperative to comprehend and implement contemporary technologies, including blockchain for secure transactions, IoT for real-time monitoring, and AI algorithms like CNN and LSTM for market trend predictions. Python was instrumental in the development of the algorithms and the integration of the platform’s various components. An interactive and user-friendly dashboard was developed using Streamlit, which enables stakeholders to visualise data and make informed decisions. In order to bridge the knowledge gap, this voyage necessitated extensive research and self-study. This enabled me to develop and execute a dependable system that predicts agricultural commodity trends and enables transparent trading.

As I progressed through the project milestones, the significance of effective time management and project planning became increasingly apparent. The adage, “If you fail to plan, then you are planning to fail,” spoke to me very deeply. I discovered the adaptability of a variety of computer vision technologies and platforms that are available on the market, including TensorFlow for the development of AI models and mm detection for object detection. Furthermore, I acquired substantial expertise in data management, which encompassed the collection, storage, and processing of extensive datasets. This was essential for guaranteeing the forecasting models’ dependability and precision. By utilising agricultural data, it was possible to gain a deeper understanding of market dynamics and the obstacles encountered by farmers and merchants. This underscored the potential of a well-designed ATP to enhance market efficiency and transparency.

Ultimately, this initiative emphasised the significance of continuous learning and interdisciplinary collaboration. Not only did my technical abilities improve, but I also gained a more profound comprehension of the dynamics of the agricultural market and the potential for digital transformation in this sector as a result of my practical experience with Python and Streamlit. The substantial influence of technology on agriculture was illustrated by the comprehensive approach, which included the integration of blockchain for secure transactions, the development of forecasting models, and data collection through IoT devices. The ATP has the potential to empower farmers by providing them with improved market access and pricing information, thereby contributing to a more efficient and transparent agricultural trading system.

7. Testing

The testing methods that we conduct automated tool testing using Selenium Python Web Driver and manual testing. Mainly testing using the test scripts which divide into function's priority whether the function have risk "High", "Medium" or "Low" as shown in **Table 2**. Next, for each different function need to create a test plan and test cases which consists of test case ID, Features ID, test script, test output, input and expected result. **Table 3** and **Table 4** show the testing environment and test cases.

1) Testing purpose

- To confirm that the ATP satisfies all prerequisites and operates as intended in a variety of situations.
- To find flaws and verify the platform's performance, dependability, and usefulness before deploying it.

2) Features to be tested

Table 2. Table of features to be tested.

Function ID	Function Description	Risk Level
F001	User Sign Up	High
F002	User Login	High
F003	View Homepage	Medium
F004	View Forecast	High
F005	View Commodity Information	Medium
F006	Add New Data in Database (Admin)	High
F007	Update Data in Database (Admin)	High
F008	Delete Data from Database (Admin)	Medium

3) Environment

Table 3. Table testing environment.

Components	Details
OS	Windows 11
Web Server	Built-in Server It starts a local server at http://localhost8501 (Streamlit) (Apache Tomcat 9.0)
Framework	Selenium WebDriver
Database	MySQL Workbench 8.0 CE
Application	ATP Web Application

4) Test cases

Table 4. Table of test cases.

Test Case ID	Description	Result
TC-00-001	Verify Sign Up with valid and invalid data	Pass
TC-00-002	Verify Login Email and Password	Pass
TC-00-003	Verify essential elements on the homepage of the Agriculture Trading Platform.	Pass
TC-00-004	Verify Forecast Pages	Pass
TC-00-005	Verify the Frequency Distribution Page	Pass
TC-00-006	Verify Add New Data in Database Page	Pass
TC-00-007	Verify 'Update Record in Database' button	Pass
TC-00-008	Verify Delete Data from Database	Pass

8. Results and discussion

The central goal of this project was to develop a dashboard that predicts the production, consumption areas, and market price patterns of Malaysia's primary agricultural commodities. The resulting web-based dashboard provides real-time insights into the agricultural sector, including map visualizations, summaries of production and consumption trends across districts and states, and live pricing updates. Feedback from evaluations indicated high usability and accessibility, underscoring its potential for wide-scale adoption. The integration of data from diverse sources facilitated precise visualizations of production and consumption trends, offering stakeholders a powerful decision-making tool.

To forecast agricultural market trends, the project utilized advanced deep learning algorithms, specifically Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN). These were compared against traditional methods like ARIMA and baseline machine learning models to demonstrate their superiority.

The LSTM model significantly outperformed ARIMA and baseline ML models, with a 23% reduction in error margins. Its ability to handle nonlinear temporal data allowed for highly accurate price forecasts. CNN excelled in identifying consumption hotspots, providing critical insights into market demand.

The integration of LSTM and CNN into the dashboard enabled real-time updates and accurate projections, empowering farmers and traders with actionable insights. These models addressed limitations in conventional techniques by capturing complex patterns and offering localized predictions. The hybrid deployment of LSTM for price trends and CNN for spatial consumption trends demonstrated the power of combining deep learning methodologies to address agricultural challenges.

Limitations and future research directions

Although the Agriculture Trading Platform (ATP) achieved success, it also faced significant restrictions during the project. The precision and dependability of the prediction models and the overall efficiency of the platform are greatly influenced by the calibre and accessibility of agricultural data. Unreliable or inadequate data might result in erroneous forecasts and understandings. Furthermore, the incorporation of various technologies such as IoT, blockchain, and AI presented considerable technical obstacles, necessitating the use of advanced technological expertise and resources to guarantee smooth communication and instantaneous data processing. Although the current platform is efficient for its first deployment, there may be difficulties in scaling the system to handle a bigger user base and more extensive data sets. This might potentially impact the system's efficiency and responsiveness. Overcoming reluctance to modifying traditional trading patterns is a problem in promoting the widespread adoption of technology among farmers and dealers, particularly those who are not technologically inclined. Additionally, the efficiency of IoT devices and the collection of real-time data depend on the presence of dependable internet and technological infrastructure, which may be insufficient in rural regions.

Subsequent investigations should prioritise the enhancement of data collection techniques, guaranteeing the excellence of the data, and incorporating a wider range of data sources to augment the platform's precision and dependability. Addressing these difficulties could be achieved by developing sophisticated IoT devices with enhanced connectivity and improved data accuracy. By delving into and deploying more advanced AI and machine learning models, it is possible to enhance market trend predictions and the mapping of consumption patterns. It is advisable to explore techniques like reinforcement learning and ensemble approaches in order to improve the accuracy of predicting. Research should prioritise the development of scalable architectures and solutions to efficiently handle greater datasets and user bases. This includes optimising cloud computing resources and utilising distributed computing approaches. In order to promote user adoption, it is important to create extensive training programmes and support systems that can help users properly utilise the platform. Potential future versions of the platform may incorporate functionalities like as automated logistics, integrated payment systems, and advanced market analytics to enhance the efficiency of trade and offer further benefits to users. Performing extensive and extended research to evaluate the socio-economic consequences of the platform on the agricultural community, such as analysing alterations in income levels, market accessibility, and overall farmers' livelihoods, will yield significant knowledge. To boost the worldwide relevance and impact of the ATP, it is important to expand the research and investigate its applicability in various geographical regions and agricultural markets. Additionally, customising the platform to address specific local concerns and market dynamics would further contribute to its effectiveness. To overcome these constraints and investigate potential areas for future study, the Agriculture Trading Platform can go further, offering enhanced advantages to the agricultural industry and promoting sustainable farming methods and economic growth.

9. Conclusion

The Agriculture Trading Platform (ATP) has the potential to revolutionize Malaysia's agricultural sector by addressing inefficiencies, improving transparency, and empowering stakeholders with data-driven insights. Several enhancements could further elevate the platform's usability, functionality, and impact:

1) Enhanced Forecasting with Advanced Models

Incorporating advanced LSTM models, such as bi-directional LSTM and LSTM with attention mechanisms, can capture more intricate time-series patterns. These improvements would provide more accurate market and production forecasts, enabling farmers to align production with market demand, minimize waste, and increase profitability. Traders and suppliers would benefit from improved logistics and inventory planning, while consumers would experience greater price stability, fostering a more resilient and efficient agricultural market.

2) Personalized and Intuitive User Experience

Customizable dashboards with widgets, notifications, and personalized reports would enhance user satisfaction and engagement. By catering to the unique needs of farmers, merchants, and other stakeholders, the platform can drive adoption and contribute to greater transparency and efficiency in Malaysia's agricultural markets.

3) Integration of Advanced IoT Sensors and Satellite Imaging

Leveraging IoT sensors and satellite imaging for real-time data collection would improve the accuracy and reliability of forecasts. Farmers could optimize crop management using precise weather and soil data, while traders and suppliers could streamline operations by understanding production cycles and market trends. This innovation would create a more informed and efficient agricultural ecosystem.

4) Automated Anomaly Detection and Data Normalization

Implementing automated systems for anomaly detection and data normalization would ensure high-quality data for analysis and forecasting. This would lead to more accurate predictions and actionable insights, enabling stakeholders to make better-informed decisions and improving supply chain efficiency.

5) Comprehensive Data Integration and Management

Developing an advanced data integration system to consolidate information from multiple sources would provide a holistic view of the agricultural market. Unified datasets would empower farmers with actionable insights into operations, while traders and suppliers could gain a deeper understanding of market trends and production levels. This would foster better coordination and a more efficient agricultural industry in Malaysia.

By incorporating these enhancements, the ATP can become a transformative tool, offering precise consumption and commodity forecasts, robust deep learning capabilities, and an improved user experience. These advancements will not only enhance the platform's utility but also contribute to a more transparent, efficient, and sustainable agricultural sector in Malaysia.

Author contributions: Conducted the research, developed the methodology, and implemented the project, MNMA; supervised the research, provided conceptual guidance, and secured funding for the project, RKR; moderated the research discussions and reviewed the manuscript draft to ensure clarity and coherence, VR. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Grant Scheme (FRGS) under Grant Code MMUE/220033, which aims to promote and facilitate innovative research initiatives in various fields of study.

Conflict of interest: The authors declare no conflict of interest.

References

Azhar, R., Kesumah, F. S. D., Ambya, A., et al. (2020). Application of short-term forecasting models for energy entity stock price (study on Indika energi Tbk, JII). International Journal of Energy Economics and Policy, 10(1), 294–301. https://doi.org/10.32479/ijeep.8715

Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos, Solitons & Fractals, 135, 109864. https://doi.org/10.1016/j.chaos.2020.109864

Department of Agriculture Malaysia. (2022). Paddy statistics. Available online: https://www.dosm.gov.my/v1/index.php?r=column/ctwoByCat&parent_id=45&menu_id=Z0VTZGU1UHBUT1VJMFlpaXRRR0xpdz09 (accessed on 1 September 2024).

Department of Statistics Malaysia. (2020). Household income and expenditure survey report 2018/2019. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=cUp6NlNndGlaQkZhK0gwYUMyWFRxdz09 (accessed on 1 September 2024).

Department of Statistics Malaysia. (2021). Gross domestic product (GDP) estimates 2021. Available online: https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=100&bul_id=ckRVWlQrNVF4K2k3M1BWYU8vVTM0Zz09&menu_id=TE5CRUZCblh4ZTZMODZIbmk2aWRRQT09 (accessed on 1 September 2024).

Department of Statistics Malaysia. (2022). Monthly External Trade Statistics. Available online: https://www.dosm.gov.my/portal-main/release-content/monthly-external-trade-statistics-apr-2023 (accessed on 1 September 2024).

Department of Veterinary Services Malaysia. (2022). Livestock statistics. Available online: https://dosm.gov.my/v1/index.php?r=column/cglossary2&menu_id=eWd2VFdIZ2xpdzBmT2Y0a0pweDcwQT09&alpha=WlFWbFFFZWw1OE9FeloxaWoxNEhrQT09 (accessed on 1 September 2024).

Guevara-Viejó, F., Valenzuela-Cobos, J. D., Vicente-Galindo, P., et al. (2021). Application of K-Means Clustering Algorithm to Commercial Parameters of Pleurotus spp. Cultivated on Representative Agricultural Wastes from Province of Guayas. Journal of Fungi, 7(7), 537. https://doi.org/10.3390/jof7070537

Malaysian Palm Oil Council. (2021). Malaysian palm oil exports set new record in 2021. Available online: https://mpoc.org.my/ (accessed on 1 September 2024).

Manogna, R. L. (2020). Innovation and firm growth in agricultural inputs industry: empirical evidence from India. Journal of Agribusiness in Developing and Emerging Economies, 11(5), 506–519. https://doi.org/10.1108/jadee-07-2020-0156

Mao, L., Huang, Y., Zhang, X., et al. (2022). ARIMA model forecasting analysis of the prices of multiple vegetables under the impact of the COVID-19. PLOS ONE, 17(7), e0271594. https://doi.org/10.1371/journal.pone.0271594

Ministry of Agriculture and Food Industry. (2022). Agriculture sector overview. Available online: https://mafi-events.com/about-ministry-of-agriculture-and-food-industries-malaysia/ (accessed on 1 September 2024).

Naorem, A., Rani, D., Roy, S., et al. (2019). Frontier soil technologies for sustainable development goals (SDGs) in India. In: Challenges and Emerging Opportunities in Indian Agriculture. National Academy of Agricultural Research Management, Hyderabad, India. pp. 113–152.

Raza, Z., Haq, I. U., Muneeb, M. (2023). Agri-4-All: A Framework for Blockchain Based Agricultural Food Supply Chains in the Era of Fourth Industrial Revolution. Available online: https://ieeexplore.ieee.org/document/10077566/ (accessed on 1 September 2024).

Sabu, K. M., & Kumar, T. K. M. (2020). Predictive analytics in Agriculture: Forecasting prices of Arecanuts in Kerala. Procedia Computer Science, 171, 699–708. https://doi.org/10.1016/j.procs.2020.04.076

Saprabha M., Dhanaraj, R. K. (2020). Machine learning based pedantic analysis of predictive algorithms in crop yield management. In: Proceedings of 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA); 5–7 November 2020; Coimbatore, India. pp. 1340-1345.

Siami-Namini, S., Tavakoli, N., Siami Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. Available online: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8614252 (accessed on 1 September 2024).

Virginia, V. E., Ginting, J., Elfaki, F. A. M. (2018). Application of GARCH model to forecast data and volatility of share price of energy: Study on Adaro Energy Tbk, LQ45. International Journal of Energy Economics and Policy, 8(3), 131-140.

Weng, Y., Wang, X., Hua, J., et al. (2019). Forecasting Horticultural Products Price Using ARIMA Model and Neural Network Based on a Large-Scale Data Set Collected by Web Crawler. IEEE Transactions on Computational Social Systems, 6(3), 547–553. https://doi.org/10.1109/tcss.2019.2914499