Predictive Modeling of Dairy Sales Using Multi-Perspective Fusion Bi-LSTM Integrated with Universal Scale CNN: Insights from the Dairy Supply Chain

[*1&2]Naveen D. Chandavarkar, [3]Dr. Soumya S
[1*]Research Scholar, Institute of Computer Science, Srinivas University, Mangalore
[2]Assistant Professor, Government First Grade College, Lingasugur, Raichur District, India
[3]Assistant Professor, Institute of Computer Science, Srinivas University, Mangalore

[1*]Corresponding Author Details: naveendchandavarkar.cet@srinivasuniversity.edu.in

Abstract

Sales prediction is a significant task of every industries. A potential prediction may majorly impact the revenue loss, out of stock and excessive stock. Many existing research have been implemented to predict the dairy sales prediction, however there are some limitations. To address this problem, the proposed research uses DL (Deep Learning) based technique to forecast the dairy sales. The proposed research uses dairy supply chain dataset to assess the proposed model CNN (Convolutional Neural Network). The present research uses Universal Scale CNN, specifically 1D-CNN, that is able to acquiring the features in ideal and in effective rates. Followed by, the extracted features are fed as an input to Multi-Perspective based Bi-LSTM (Bidirectional Long Short Term Memory) that is able to acquiring the features in an effective manner in characteristics of reducing the error rates upon the prediction sales rate of dairy based products. The performance of proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN is evaluated by different performance metrics which includes RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), MSE (Mean Square Error) and R^2 (R Square). When compare to other models, the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN outperforms with optimal performance value with high R^2 value of 0.9824. The performance metrics provides broad analysis in terms of accurate prediction of the proposed model.

Keywords: Deep Learning (DL), Convolutional Neural Network (CNN), Bidirectional Long Short Term Memory (Bi-LSTM), Dairy Sales Prediction

1. Introduction

One of the most broadly used products in each community is dairy products. The dairy products are extremely considered due to its high benefits and properties. The necessity for all types of dairy products is very capricious. However, according to medical advice, every individual need to consume more dairy products [1]. On the other side, provided the competitiveness of this industry, the quality and price of the product have a major impact on increasing or decreasing demand. Due to this reason, the demand behaviour for dairy products is more complicated [2]. In these environment, predicting the upcoming demand for dairy products is more significant for any of the industries in this sector. Identifying the precise prediction of the PDP is something which company executives are constantly concerning [3]. However, because of its high complication of product demand, common techniques such as regression might not possible. Due to this reason, intelligence techniques are required to predict DPD. Some existing research recommended to employ NN (Neural Network) and SVR (Support Vector Regression) [4]. The predictive models are relevant to cow behaviours, milk quality, health and productivity are the significant components of the recent dairy industries, as these representations can aid with decision making in farm related product marketing [5]. For an instance, forecast of DMY (Daily Milk Yield) is esteemed for classifying cows which have failed to obtain their forecasted path in milk yield for causes of welfare or other disease problems. It may also aid in choosing dairy cows earlier for breeding programmes and genetic improvements [6]. The lactation bends can aid farmers to pinpoint cows underneath negative energy balance and recognize animals which are more resilient to metabolic stress so that they may improve the feeding process, comprising pointing cows which are losing too much weight by providing them with high quality supplements like bypass protein [7]. The forecast of milk composition offers significant information nearby the predicted purchase price of milk that may assist farmers to take satisfactory business decisions. Additionally, the forecasting of MF (Milking Frequency) is another significant part of information which might be helpful in calculating cow welfare and management efficacy in relation to enhancing the traffic system management in RMS. Where, regression models are the common way of generating models with cow relevant factors. The essential algorithms of regression models significantly contains both traditional statistics and ML (Machine Learning) [8].

The earlier category of algorithms based on the techniques of standard linear regression such as generalized linear regression, stepwise multiple linear regression and segmented multi-phases regression. Conversely, the real-time application and forecast potential of such linear regressions may be firmly restricted because of their structural restrictions and quality of the available data [9]. In recent times, ML algorithms are progressively recommended by many companies pursuing to build real time applications, because of their better accuracy and enhanced flexibility. Thus far, forecasting models have been implemented for the dairy industries by SVR (Support Vector Regression), NN regression and RF (Random Forest) [10]. In recent decades, in many countries major amount of data produced by RMS are available to model developers. Deployed forecasting models majorly focused on estimate at a herd level which means 305 day milk yield [11]. The upgraded enhanced forecast models are estimated to be able to forecast features which are relevant to individual cow relatively than the entire kettle and over extended time periods, to permit sudden and accurate problem solving on farms. This encompasses the recognition of separate requirement that in turn minimizes the probabilities of long term loss in productivity [12]. ML models such as RF can hold classify data and unresponsive to missing values. Moreover, they have the potential to analyze huge datasets that frequently complicate to compute with conventional statistical models. This emphasis the estimation which ML techniques might provide for dairy farming. Evaluating huge incorporated datasets can permit offering farmers with satisfactory decisions support systems such as help them to maximize the efficacy and well-being of their animals [13]. Enhancements in sensing technologies which have led to maximized obtainability of sensors in farming. Milking machines that provide data of daily milk yield are the utmost general deployed sensors in dairy farming. The systems which observers separate animal behavior such as rumen pH, rumination, location and estrus are also become available. In addition, many farmers save data on the concentrate feeding behavior of separate cow [14]. The ML models, between other techniques which represent a method to recognizing these datasets that are explicitly available on many farms. ML aims to efficiently reproduce the behaviour of human learning, permitting for the automatic acquisition and detection of new information. The algorithms are able to identify clusters in larger datasets with huge variables, forecasting the events and learn from the fed data [15]. ML is already employed in the dairy sector in quite diverse ways. While, there is a limitations of overview of the algorithms employed. Challenges in implementing the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model in the dairy sector include ensuring data quality and availability, managing the complexity of advanced algorithms [13], and requiring substantial computational resources. Additionally, integrating the model with existing systems, ensuring interpretability of results, and overcoming resistance to change among stakeholders can hinder effective adoption and utilization [15]. To address this problem, the proposed research employs Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model to predict the dairy sales which aid farmers to plan accordingly. The proposed research employs dairy supply chain dataset to evaluate the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN. The performance of the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model is evaluated by MSE, MAE and RMSE.

1.1 Research Contribution

The main objective of the proposed research as follows:

- To employ dairy supply chain dataset for predict the sales of dairy products.
- To demonstrate enhanced regression model which is proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model for accurate prediction of dairy product sales with minimal error rates.
- To implement proposed Universal Scale CNN model to monitor and conceal all the corresponding scale field by transforming the input between innumerable associations.
- To analyze the performance of the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN with the performance metrics.

1.2 Paper organization

The framework of the research paper is organized with overview in Section I. The review of prevailing studied involved in dairy products sales prediction are described in Section II along with research gaps. Section 3 represents the proposed methodology and the process involved. Section IV deliberates the results achieved by the proposed model. Section V deliberates the conclusion and future work of the proposed research.

2. Literature Review

The Predictive Modeling of Dairy Sales using ML and DL algorithms are discussed in the following section. Correspondingly, the study [16] has compared the performance of two demand forecasting models, ARIMA and LSTM (Long Short-Term Memory), by utilizing 5 years of data from 8 dairy products across 5 production plants. Both the models are found to be reliable, with ARIMA performing better on less fluctuating data and LSTM excelling in capturing seasonal trends and strong patterns. As a result, training on monthly data yields better error scores. As several challenges has been faced in the prediction of milk yield in real time operation, the prevailing study [17] has suggested the GB(Gradient Boosting) algorithm for predicting milk yield on data from the top three largest farms sourced from an AgTech platform, Farm trace. Finally, the outcomes has implied that the maximum deviation in predictions is 2.2% for one of the farms (Farm II) and has effectively enhanced the milk yield. Similarly, in order to overcome the issues of genetic profiles, milk production records and to predict the milk production at both short and long term levels, the study [18] has utilized the conventional ML models and innovative DL algorithms. As a result, the suggested models has enhanced the dairy operations by identifying key patterns, optimizing resource allocation, and enhancing bovine health management which allows the ability to the farmers for adapting quickly to environmental changes, manage risks more effectively, and increase productivity. Alternatively, the study [19] has revealed the challenges faced by the prevailing studies as data complexity, inconsistencies across herds, and limited variable scope. In order to tackle the challenges, the study has utilized the decision trees for identifying the key factors affecting milk yield in Polish Holstein-Friesian cows. As a result, the study has found that the cows milked more than three times a day has the highest yield of 47.24 kg, while the others has the lowest yield of 13.56 kg. Similarly, the study [20] has suggested several variable reduction methods such as PCA (Principal Component Analysis) and DT, for overcoming the issues of overfitting in prediction models from two branded milks from 2007-2008 by the 9 storage stations in Ishikawa Prefecture. Though the above mentioned methods has been utilized in the phase-1, the phase-II has utilized other methods such as KNN (K- Nearest Neighbors), RBF (Radial Basis Function) for assessing the presentations. As a result, the KNN algorithms has obtained better outcomes in terms of accuracy in sales estimation. Correspondingly, the prevailing study [21] has utilized the ML methods with a 5 year data which is based on the features, production and health monitoring of 80 cows in the robotic dairy farm. Additionally, the ML methods has been utilized to overcome the issues faced by the existing studies and to enhance the efficacy of the RMS (Robotic Milking Systems). Also, the time series cross-validation has been implemented for evaluating the presentation of the model under certain situations which has provided increased accuracy rates and the ML methods has the ability of enhancing the farm control and animal welfare in the farms.

Similarly, the study [22] has framed a hybrid model in the combination of RF, SVR and k-medoids for forecasting the curves of lactation utilizing the monthly milk yield and also to tackle the issues of lactation curves. Additionally, the framework has been implemented on the historical milk production information by 260 Wisconsin dairy farms between 2010-2014. As a result, the K-R-S model has obtained average outcomes with a decrease in MAE of 24.2% and 23.4% of RMSE and has increased correlation. Accordingly, the existing study [23] has contrasted both the ANN (Artificial Neural Networks) and DA for forecasting the milk yield level of 3460 cows in a Holstein Friesian cattle in 16 years. As a result, the ANN has obtained better outcomes in terms enhanced accuracy than DA. Additionally, the AUROC (Areas under Receiver Operating Characteristic) curves has been involved and has provided enhanced demonstration in ANN depending on various calving periods. Correspondingly, the prevailing study [24] has utilized the RF for forecasting the milk yield regarding the environmental attributes. Additionally, the framework has been created and assessed by utilizing the 91 cows which are lactating in a northern Italy dairy form by implementing the THI (Temperature – humidity Index) and the lactation phase for recording the outcomes of the heat stress. As a result, the introduced RF has obtained an error of 18% and 2% of milk yield which has implied that the ML has the ability to forecast the milk yield by enhancing yielding and efficacy. Alternatively, the study [25] has suggested several mathematical operations which involves the wilmink and wood models for locating the lactation curves and also for tackling the issues of accurately modifying the lactation curves of the dairy cows particularly during the lactation phase. Moreover, the involved algorithms has been evaluated for their capability for predicting the delayed milk production in terms of robustness and precision. As a result, the models has obtained average level of prediction accuracies in various situations. Subsequently, the study [26] has utilized the CNN for examining the images with ultrasound for forecasting the milk production and the lactation phase. As a result, the DL algorithms has obtained better outcomes in terms of accuracy forecasting in contrast to the prevailing studies. In addition to this, the study [27] has utilized several forecasting techniques for assessing the FL305DMY (First Lactation 305-Day Milk Yield) by utilizing the 2100 double monthly tested milk production data of the 350 Murrah buffaloes such as CDM (Centering Date Method), RM (Radio Method), TIM (Test Interval Method), ANN and MLR (Multiple Linear Regression). As a result, the MLR has obtained average outcomes in terms of accuracy along with a decrease in RMSE and increase in R^2 in which the whole 6 days data has been utilized.

Correspondingly, the study [28] has introduced a technique by utilizing the CV (Computer Vision) along with the walk-over measuring scales for evaluating the milk production by a single cow and the presentation of the lactation and has involved the 3D imaging of the udder and its corresponding weight before and after the production of milk. As a result, the introduced algorithms has obtained average outcomes in terms of R^2 and RMSE. In addition with the model has deliberated the ability of trait analysis like length and teat placement with an average level of accuracy. Alternatively, the study [29] has implemented the RF algorithm for forecasting the daily milk production regarding the situations of the surroundings especially in the heat stressed area which has been recorded by the THI (Temperature- Humidity Index) and to tackle the issues of using the numerous of real time data. As a result, the suggested model has obtained average outcomes in terms of accuracy and detection error in the daily milk production of a single cow. Correspondingly, the existing study [1] has recommended a RMP-CPR (Raw Milk Price Prediction Framework) for evaluating the presentation of the consumer concerning the relationship of milk value and dairy ingesting. As a result, the suggested model has improved the growth of precise marketing plans and has a main role in corresponding the interests among consumers. Alternatively, the study [30] has utilized certain algorithms such as RF ,GB , SVM , NN , and AdaBoost, for predicting the price of the milk. Additionally, the ML method has involved the HPT (Hyper Parameter Tuning) along with a nested cross-validation for considering the strength of the introduced algorithms. As a result, the introduced algorithms has attained better outcomes in terms of accuracy, precision, F1 score and recall. Subsequently, the existing study [31] has introduced a SAE (Sequential Auto Encoder) for acquiring the abstract and interpret the low-dimensional pictures and the sequential data . As a result, the hybrid lactation number along with the herd statistics followed by the day to day milk production which the cow has overcome in the period of lactation has been exhibited in the qualitative dormant depictions. Correspondingly, the study [32] has recommended the LSTM which is based on the NNs for enhancing the decision. Additionally, the data which has been gained from 6000 diverse herd. As a result, the recommended method has achieved enhanced outcomes in terms of verification time, by predicting the regular income for every cow's 5th year with a RMSE. This study explores the modernization of the dairy industry, focusing on enhancing dairy supply chain (DSC) optimization through machine learning and artificial intelligence. By analyzing existing strategies and emphasizing waste reduction and quality improvement, the research aims to boost operational efficiency while addressing environmental concerns and economic challenges faced by smaller farms. The findings highlight the shift from traditional mathematical modeling to advanced AI/ML methods, providing valuable insights for farmers and distributors to optimize dairy processing facilities effectively [33]. This study introduces a predictive analytics model that harnesses artificial intelligence to effectively manage food supply chains amidst fluctuating consumer preferences and global disruptions. By utilizing real-time data from IoT devices across various sources, the model enhances visibility and enables precise demand forecasting and inventory optimization, thereby reducing waste and improving resource allocation. Emphasizing adaptive strategies, it fosters collaboration among supply chain partners, promoting resilience and efficiency while contributing to sustainability and food security on a broader scale [34]. This study highlights the transformative impact of predictive analytics powered by machine learning and artificial intelligence in the dairy industry. By leveraging extensive datasets, including milk production records and environmental data, the research develops decision support systems that forecast milk production and identify key influencing factors. Utilizing individual sensors on approximately 4,000 cattle from a major Jordanian dairy farm, the models predict both short-term and long-term milk yields, enabling farmers to make informed decisions, improve resource allocation, and enhance overall productivity and bovine health care practices [35]. Increasing the economic and environmental sustainability of dairy cattle hinges on accurately forecasting milk production throughout a cow's career, which is crucial for efficient management. While many researchers focus on selecting high-yield animals, there is a notable gap in tools that predict future lactation productivity based on historical data. Understanding a cow's potential for high, medium, or low milk output, especially during its first few years, is vital for farmers. The advent of large datasets from automatic milking systems allows for the application of machine learning algorithms, particularly classification learners, to enhance predictive capabilities in this area [36]. This study aimed to develop computational models for predicting the quality of bulk-tank milk in dairy sheep and goat farms by utilizing health management-related variables. By analyzing milk from 325 sheep and 119 goat farms, five target values—fat content, protein content, combined fat and protein, somatic cell counts, and total bacterial counts—were predicted using various machine learning tools. K-nearest neighbors emerged as the most effective model for predicting protein content in both sheep and goats, while neural networks and random forests were also successful for specific metrics. The results provide valuable insights for farmers and dairy processing factories, facilitating rapid assessments of milk quality and enabling targeted interventions to enhance production standards [37]. This study proposes a framework integrating blockchain, IoT, and cloud technologies to enhance visibility and reduce costs in the Australian milk supply chain. Utilizing a design science research methodology, the findings indicate that this integration yields a high return on investment for stakeholders, with varying benefits across the supply chain. The research aids milk industries in making informed decisions about technology adoption, ultimately improving trust and coordination among stakeholders and consumers [38]. The paper proposes a blockchain-based layered architecture to enhance traceability and information sharing in dairy product supply chains, addressing challenges such as data loss, quality inconsistencies, and single points of failure. By leveraging sharding for parallel processing, the architecture supports various components, including quality assurance, automated payment settlements, sales analysis, and consumption monitoring. Utilizing Hyperledger Fabric, the study details the algorithms for traceability chain codes and offers a comparative analysis demonstrating the advantages of Hyperledger over other existing systems, highlighting its suitability for supply chain networks [39]. The paper emphasizes the critical need for an effective traceability system in the dairy industry, particularly for perishable products vulnerable to fraud, adulteration, and safety concerns. It outlines how such a system can enhance visibility and transparency, promote consumer trust, optimize operations, and support sustainable development goals by minimizing waste and improving resource utilization. Through content analysis, the study investigates the theoretical and technological landscape of dairy traceability, exploring the role of digital technologies and identifying research opportunities. It discusses the evolution of traceability, current challenges, and advancements, highlighting its implications for stakeholders in achieving specific United Nations Sustainable Development Goals (UN SDGs) [40]. This paper presents a comprehensive framework for integrating blockchain technology (BCT) into agri-food supply chains (AFSC), focusing on a case study involving honey and coriander powder. Collaborating with an agri-food organization and a technology provider, the authors developed a proof-of-concept (POC) for BC-enabled traceability. The study highlights that the increasing demand for improved traceability can drive adoption, despite existing challenges in implementation. A cost framework is introduced to analyze developmental and operational expenses associated with the POC. Additionally, the research explores stakeholders' perspectives on implementation challenges and perceived benefits, emphasizing BCT's potential at the intersection of digitalization and product integrity. This study serves as a practical guide for managers, providing insights into the operational and strategic aspects of adopting BCT in AFSC [41]. This study investigates the implementation of blockchain (BC) technology in food traceability, particularly focusing on protected designation of origin feta cheese production in Greece. It conducts a comprehensive assessment encompassing technical, environmental, and economic dimensions using both private (Quorum) and public (Ethereum) blockchain systems. The findings reveal that the BC-based system has minimal environmental impacts compared to traditional feta production methods. Economically, the system proves beneficial for dairy producers, especially when accounting for consumers' willingness to pay a premium for BC-traceable products. Sensitivity analyses indicate that scaling up—from increasing transaction volumes to expanding production—can enhance the advantages of BC-based traceability, highlighting its potential for wider application within the dairy industry [42]. The COVID-19 outbreak has exposed vulnerabilities in food supply chains (FSCs), highlighting the urgent need for enhanced resilience due to the perishable nature of food products. This article identifies key enablers of resilience through a review of existing literature and employs a graph theory matrix approach to analyze their interrelationships and importance. The findings prioritize enablers such as readiness, stakeholder collaboration, IT alignment, risk awareness, responsiveness, flexibility, appearance, and sustainability, offering valuable insights for policymakers and managers in strengthening FSCs against future disruptions [43]. This study analyzes retail sales prediction for Walmart by combining machine learning and time series models to enhance forecasting accuracy amid complex consumer behaviors. Utilizing various methodologies, including ARIMA and Prophet, and incorporating external factors, the findings reveal that hybrid approaches significantly outperform traditional models, offering valuable insights for optimizing inventory management and customer satisfaction [44]. This paper explores enhancing demand forecasting accuracy in retail supply chains by applying machine learning techniques, specifically ARIMA and XGBoost models, to address the limitations of traditional forecasting methods. While ARIMA effectively captures seasonal trends, XGBoost excels in modeling complex relationships, such as fuel prices and CPI. The combination of both models leads to significant improvements in forecast accuracy, suggesting potential for more efficient inventory management strategies in the retail sector [45].

2.1 Problem Identification

- The existing study [16] has the challenges in integrating the real-time data from diverse sources for improving the accuracy of demand forecasting models for dairy production, particularly under varying environmental conditions.
- The study [26] has included major drawbacks such as enhancing the robustness of ultrasound-based deep learning models for milk yield prediction, particularly in diverse farming environments with varying cow health and management practices.
- The limitations of the study [29] includes optimizing the RF algorithm for better generalization across different climates and farm conditions, as well as incorporating more granular environmental and biological data to improve predictions under heat stress.

3. Research Methodology

3.1 Proposed Method

In dairy industries, sales forecasting is one of the primary task in recent times which aids farmer to take decisions based on the predictions. The proposed research employs CNN based model to predict dairy sales using publicly available dataset. The overall process of proposed model is represent in the figure 1.

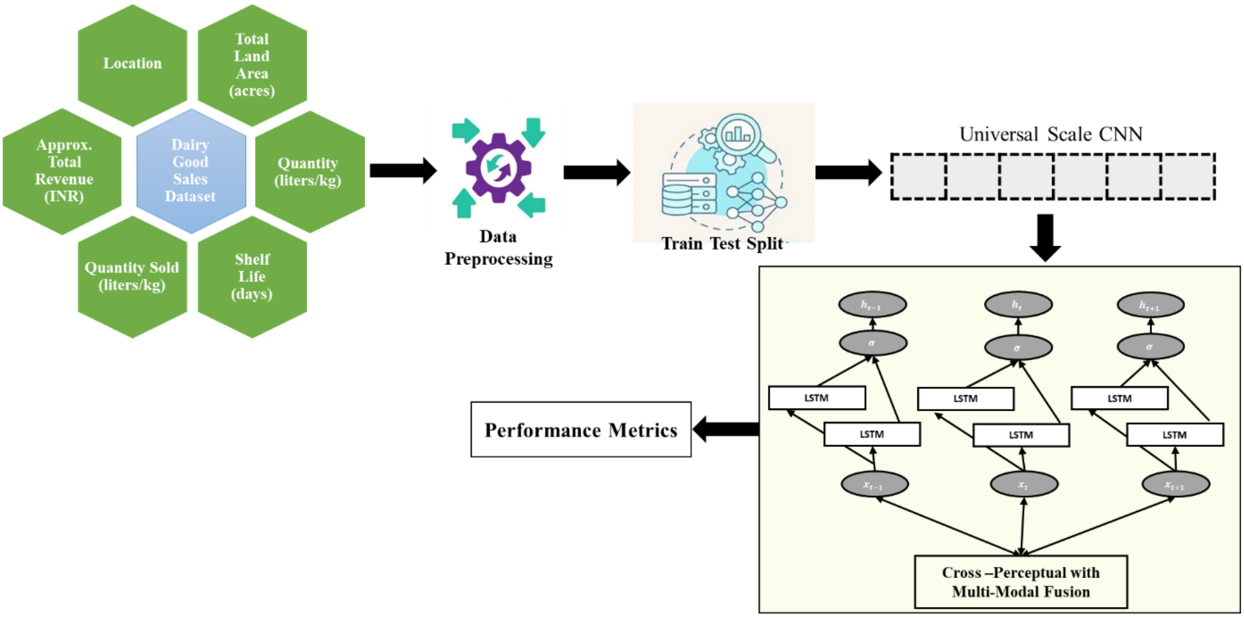

Figure 1. Overall Process of Proposed Model

Figure 1 depicts the overall process of proposed research. The proposed research employs dairy supply chain dataset to evaluate the proposed model. The dataset encompasses several attributes for prediction such as geographical data of dairy sales, production ability (total land area). Sales quantity, total income, aggregate available quantity for sale and time of the dairy products remains sellable. The dairy sales dataset undergoes pre-processing stage to enhance the process of the proposed model by convert the missing values into null, remove outliers and label encoding. The processed data is split into train (80%) and test (20%) dataset. The train data is processed by the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN. The proposed research evaluate the efficacy of the model by test data. The performance of the proposed research is evaluated by the performance metrics of MSE, MAE and RMSE.

3.2 Dataset Description

The proposed research employs the dairy supply chain dataset to evaluate the performance of proposed model and the dataset link is provided for the reference https://test.ieee-dataport.org/documents/dairy-supply-chain-sales-dataset. The dataset contains data related to daily sales of a series of a dairy product codes provided by MEVGAL The dataset includes the sales information collected from logistics division from 2020 to 2022. The features of the dataset are day, month, year, daily_unit_sales, previous_year_daily_unit_sales, percentage_difference_daily_unit_sales, daily_unit_sales_kg, previous_year_daily_unit_sales_kg, percentage_difference_daily_unit_sales_kg, daily_unit_returns_kg, previous_year_daily_unit_returns_kg, points_of_distribution and previous_year_points_of_distribution.

3.3 Data Pre-processing

3.3.1 Feature Scaling

Feature scaling is significant to confirm which all characteristics subsidize evenly to the evaluation, especially in algorithms which are delicate to the data scale. Where normalization and standardization takes place. The normalization rescales the characteristics to a usual range generally (0, 1). It is especially aided when characteristics have dissimilar scales or units. The standardization is employs to convert the characteristics to have mean and standard deviation of 0 and 1. It is useful when the data adhere to Gaussian distribution which permitting for ideal performance of ML models. The scaling confirms the metrics like sales pricing and volumes do not exclusively impact model training and forecasting because of their inherent dissimilarities in magnitude.

3.3.2 Label Encoding

It is employed to transform the categorical variables to numeric format that is require for ML algorithms which need numerical input. Each exclusive category value is allocated an integer. This conversion permits the proposed algorithms to understand categorical data efficiently. The label encoding may aid in modeling features such as supplier categories or product type and simplifying ideal visions into sales patterns.

3.3.3 Outlier Removal

This process is significant for improving the effectiveness of statistical analysis and ML models through excluding maximum values which could skew outcomes. This confirms that sudden drops and spikes in dairy supply chain sales data probably because of the rare events or error and it avoid misinform decision making process considering demand predicting.

3.4 Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN

To tackle the problems in the existing dairy sales forecasting, the proposed model employs proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN for precise prediction. The entire process of proposed model is represented in the figure 2.

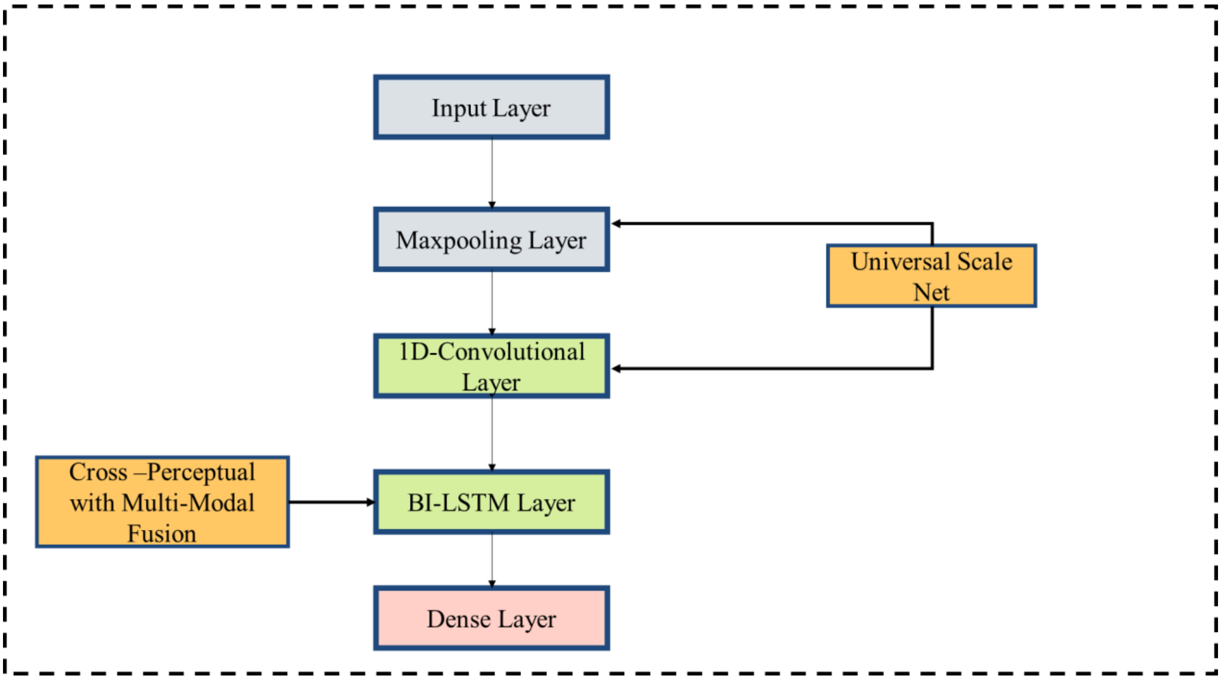

Figure 2. Entire Process of Proposed Model

Figure 2 depicts the entire process of proposed model. The process begins with the input layer, it process the input features of dairy supply chain sales dataset and fed as an input to the maxplooling layer. This layer minimizes the dimensionality of the input data however maintaining important features and makes the proposed model become computationally potential. Then, the 1D convolutional layer employs filters over the temporal data to identify significant features, trends and patterns. Here, the proposed research integrate universal scale net in maxplooling and 1D convolutional layers, this improves their robustness in dairy sales forecasting by enhancing feature extraction, maximizing computational efficacy and minimizing information loss. This improvements led to more precise forecast which permitting dairy industries to take decisions on the basis of consistent predictions resultant from their data. The proposed research enhances forecast by integrate the cross perceptual multi-modal fusion in the Bi-LSTM layer which majorly improves the forecasting by enhancing learning dynamics, feature representation, predictive accuracy and robustness of the proposed model. This permits a widespread insights of the factors which impacts sales eventually lead to optimal decision making in the dairy industries. Finally, the dense layer combined representations to take the final forecast. Where, the provided time series ϱ, using a neural convolutional process within the kernel γ is parallel to employing signal processing convolutional with kernel ψ that is a preserved vector of γ. Furthermore, in the Fourier domain, this process is also parallel to the element-wise proliferation among ϱ and ψ.

$$\varrho * \gamma = \varrho \otimes \psi = \mathcal{F}^{-1}\big(\mathcal{F}(\varrho) \cdot \mathcal{F}(\psi)\big) \tag{1}$$

The convolutional signal processing of ϱ and ψ is,

$$(\boldsymbol{\varrho} \otimes \psi)[\mathrm{P}] = \sum_{r=0}^{R-1} \boldsymbol{\varrho}[P - \mathrm{Q}]\psi[\mathrm{Q}] \tag{2}$$

Where, $\otimes$ indicates a convolution of signal processing and the convolutional theorem (1) expresses,

$$\boldsymbol{\varrho} \otimes \psi = \mathcal{F}^{-1}\big(\mathcal{F}(\boldsymbol{\varrho}) \cdot \mathcal{F}(\psi)\big) \tag{3}$$

Where, the Fourier transform is indicates by $\mathcal{F}$ and the inverse Fourier transform is indicates by $\mathcal{F}^{-1}$. In a traditional 1D-CNN, the convolution of $\boldsymbol{\varrho}$ and γ is described as:

$$(\boldsymbol{\varrho} * \gamma)[\mathrm{P}] = \sum_{r=0}^{R-1} \boldsymbol{\varrho}[\mathrm{P} - \mathrm{Q}]\gamma[\mathrm{R} - 1 - \mathrm{Q}] \tag{4}$$

Whereas, the neural convolution is indicated by $*$. While $\psi[\mathrm{m}] = \gamma[\mathrm{R} - 1 - \mathrm{Q}]$, the proposed research have the equation 1 which is acknowledged in the neural convolutional process. $\psi[\mathrm{m}] = \gamma[\mathrm{R} - 1 - \mathrm{Q}]$ signifies ψ has the preserved sequence of γ for an instance $\gamma = [1,2,3]$ at that time $\psi = [3,2,1]$. Throughout the process of time series $\boldsymbol{\varrho}$, if the 1D-CNN architecture may minimize the noise, then the extracted depiction from the process of neural convolutions with kernel γ will include minimal noise. The enhanced depiction may cause advantages to the classification. In the frequency domain, if the proposed research examine the time series, the proposed research decompose a time series data of γ into three partitions as follows.

$$\mathcal{F}(\boldsymbol{\varrho}) = \mathcal{F}(\boldsymbol{\varrho}_s) + \mathcal{F}(\boldsymbol{\varrho}_{N_s}) + \mathcal{F}(\boldsymbol{\varrho}_{N_o}) \tag{5}$$

Whereas, $\boldsymbol{\varrho}_s$ indicates the preferred signal for classification task, $\boldsymbol{\varrho}_{N_s}$ indicates the noise on the similar frequency of $\boldsymbol{\varrho}_s$ and $\boldsymbol{\varrho}_{N_o}$ indicates the noise on the remaining frequencies. In the similar way, γ the neural convolutional kernel may be decomposed,

$$\mathcal{F}(\psi) = \mathcal{F}(\psi_s) + \mathcal{F}\big(\psi_{N_s}\big) + \mathcal{F}\big(\psi_{N_o}\big) \tag{6}$$

To streamline, let's substitute $\mathcal{F}(\boldsymbol{\varrho}_s), \mathcal{F}\big(\boldsymbol{\varrho}_{N_s}\big), \mathcal{F}(\boldsymbol{\varrho}_{N_o})$ with $\mathrm{Q}, \mathrm{N}, \mathfrak{x}$ and interchange $\mathcal{F}(\psi_s), \mathcal{F}\big(\psi_{N_s}\big), \mathcal{F}\big(\psi_{N_o}\big)$ with α, β, ξ. Hence, in the frequency domain the convolutional result is

$$\mathcal{F}(\boldsymbol{\varrho} * \gamma) = \mathcal{F}(\boldsymbol{\varrho}) \cdot \mathcal{F}(\psi) = \begin{array}{c} \mathrm{Q}\alpha + \mathrm{Q}\beta + \mathrm{Q}\xi + \\ \mathrm{N}\alpha + \mathrm{N}\beta + \mathrm{N}\xi + \\ \mathfrak{x}\alpha + \mathfrak{x}\beta + \mathfrak{x}\xi \end{array} \tag{7}$$

Where, $\mathcal{F}(\boldsymbol{\varrho}_{N_o})$ is not interconnect with $\mathcal{F}(\psi_s)$ in the Fourier domain, hence the result is,

$$\mathcal{F}(\boldsymbol{\varrho}_{N_o}) \cdot \mathcal{F}(\psi_s) == 0 \tag{8}$$

In the Equation (8), for similar reason, the significance of $\mathrm{Q}\xi, \mathrm{N}\xi, \mathfrak{x}\alpha, \mathfrak{x}\beta$ are all zero, hence the result of convolution may also be represented as

$$\mathcal{F}(\boldsymbol{\varrho} * \gamma) \quad = \quad \begin{array}{c} \mathrm{Q}\alpha + \mathrm{Q}\beta + 0 + \\ \mathrm{N}\alpha + \mathrm{N}\beta + 0 + \\ 0 + 0 + \mathfrak{x}\xi \end{array} = (\mathrm{Q} + \mathrm{N})(\alpha + \beta) + \mathfrak{x}\xi \tag{9}$$

Thus, if the proposed model may minimize or delete the noisy part $\mathfrak{x}\xi$, the convolution process outcome is $\mathcal{F}(\boldsymbol{\varrho} * \gamma)$ will have minimal noise that might advantage for the classification. In some proposition, the $\mathfrak{x}\xi$ noise may not be distant by neural operators which contains Bias, ReLU and Batch norm. To be detailed, the proposed research represent the full equation of the layer result before the activation layer:

$$LayerOutput = BN(\boldsymbol{\varrho} * \gamma + bias) \tag{10}$$

Whereas, BN signifies batch norm. If the proposed research develop BN, then Equation (10) will be,

$$r(\boldsymbol{\varrho} * \gamma + bias - E) + \beta \tag{11}$$

Hence, the parameters of the Batch norm is indicated by r and β, in a batch, the mean value of $\boldsymbol{\varrho} * \gamma$ is represented by E. Thus, r, $bias, E$ and β are the trainable parameters, the Equation 11 has the similar analytic form.

$$LayerOutput = r(\boldsymbol{\varrho} * \gamma + \epsilon)$$

Whereas, the zoom rate of representation is denoted as r, $\epsilon = bias - E + (\beta/\gamma)$ is an inductive bias which is evaluates using four numeric variables which do not comprise any frequency data.

$$\mathcal{F}(LayerOutput) = r[\mathcal{F}(\boldsymbol{\varrho} * \gamma) + \mathcal{F}(\epsilon)] \tag{12}$$

Integrating Equation (9) into Equation 12, the result is,

$$\mathcal{F}(LayerOutput) = r[(S + \mathrm{N})(u + \beta) + \mathfrak{x}\xi + \mathcal{F}(\epsilon)]$$

Commencing the explanation of Batch norm and Bias, ϵ is an actual value vector of constant value. The proposed research represents the constant value as $Total\_bias$, the outcome is,

$$\mathcal{F}(\epsilon)[\mathrm{P}] = \begin{cases} Total\_bias = 0 \\ 0\mathrm{P} \neq 0 \end{cases}$$

Hence, $\mathfrak{x}\xi$ signifies the complex value vector, however $\mathcal{F}(\epsilon)$ is an actual value vector, the $\mathcal{F}(\epsilon)$ noise may not be represented by ϵ in neural convolution. The kernel sizes are a prime number from 1 to *N*, in the initial two convolutional layers and each kernel sizes are dissimilar to each other in the layer. The receptive fields of the main two layers may be any positive odd numeric from *0* and *2N*, according to the Goldbach Conjecture. In the 3rd convolutional layer, it conduct a convolutional process and it only has kernels of size 1 and 2. With this circumstances, the receptive fields of OS-CNN may employs all probable integer from *0* and *2N*, moderately than only odd numbers. If the stride is 1, the size of the receptive field is evaluated as:

$$R_{size} = 1 - L + \sum_{l=1}^{L} \mathrm{R}_l \tag{13}$$

Whereas, the total count of convolutional layers are represented as L and the kernel size of the lth layer is denoted as R_l. In difference to traditional CNN, the proposed research employs a global average pooling once the addition of one pooling layer for each convolutional layer. By employing the global average pooling in 1D-CNN which permits the operation of the CAM (Class Activation Map) to recognize the data region subsidizing to the particular labels. The CAM is evaluated as,

$$C_i = \frac{1}{\mathrm{P}} \sum_{\mathrm{P}=1}^{\mathrm{R}} F_{i,\mathrm{P}} \tag{14}$$

Whereas, the output value of the ith channel is denoted as C_i, the length of the time series is represented as P and the Pth's value of the ith channel is denoted by $F_{i,\mathrm{P}}$. The predicted output c is evaluated by the provided input $\boldsymbol{\varrho}$,

$$c = \underset{c}{argmax} \frac{1}{\mathrm{R}} \sum_{\mathrm{Q}=1}^{\mathrm{R}} P(c|\theta_m, \boldsymbol{\varrho}) * I[c = F(\theta_m, \boldsymbol{\varrho})] \tag{15}$$

Whereas, the model parameters of the mth classifier is denoted as θ_m, the possibility of the predicted class c is $P(c|\theta_m, \boldsymbol{\varrho})$ that $\boldsymbol{\varrho}$ depends and I denotes the Boolean indicator for class c is the resultant forecast of the mth classifier. Through interchanging kernels with receptive fields, the proposed research may minimize the model size and the representation capability of the proposed model will not lose. There is no sign of representation capability loss in model A. For $\forall_{\boldsymbol{\varrho}}$ and θ_{N}, $\exists\theta_{\mathrm{Q}}$ makes $\mathrm{Q}(\boldsymbol{\varrho}, \theta_{\mathrm{Q}}) = N(\boldsymbol{\varrho}, \theta_{\mathrm{N}})$. Here, the input data is denoted as $\boldsymbol{\varrho}$, the parameter of the proposed model is denoted as θ_*, the result of the proposed model parameter with output data is $*(\boldsymbol{\varrho}, \theta_*)$. The proposed research states: $L(C) = L(C_1) + L(C_2) - 1, C(K) = 1, C(\mathfrak{x}_1) = min(\mathfrak{x}_1, \mathfrak{x}_2)$, which makes,

$$Conv(X, \mathfrak{x}) = Conv(Conv(X, \mathfrak{x}_1), \mathfrak{x}_2)$$

Whereas, the input signal is denoted by X, the kernel of the one layer network is $\mathfrak{x}$, the 1st and 2nd kernels of the two layer network is denoted by $\mathfrak{x}_1$ and $\mathfrak{x}_2$. The kernel size of $\mathfrak{x}$ is denoted by $L(\mathfrak{x})$, the output channel number of K is denoted as $C(\mathfrak{x})$ and the convolution results of X and $\mathfrak{x}$ with ReLU and Batch norm is denoted by $Conv(X, \mathfrak{x})$. The multimodal features obtained as a result prominently consists of massive amount of concluded information, if the communication of modality data is not considered, which is improper for the prediction task. Hence, the proposed research utilizes the cross-perceptual with the multi-modal fusion into the fusion function called Bi-LSTM. Also, the cross perceptual feature matrices are assessed for obtaining the information matrix where the attention distribution is assess using a line by line process of softmax function.

4. Results

This section deliberates the outcomes of the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model with the performance metrics, EDA (Exploratory Data Analysis) and performance analysis of the proposed research. The demonstration of the proposed research is evaluated based on RMSE, MAE and MSE.

4.1 Performance Metrics

4.1.1 RMSE

Root Mean Squared Error is considered as the standard deviation of variance between residuals such as actual and predicted values. Lesser RMSE values signifies that the data fits well whereas high RMSE values recommend with less precise predictions and greater errors. RMSE is calculated using equation (16)

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{N}(\text{ acutal} - \text{predicted})^2 \;\; /\mathrm{N}} \tag{16}$$

4.1.2 MAE

MAE assists as an efficient metric for assessing the accuracy of proposed model, providing a balance among effectiveness and simplicity in performance evaluation.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{17}$$

4.1.3 MSE

It is the measurement of image excellence metric. If the standards are nearer to zero, the metric dimension have better quality. The formula for MSE is mentioned in equation (18)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\mathrm{Y_i} - \mathrm{Y_i})^2 \tag{18}$$

4.2 EDA (Exploratory Data Analysis)

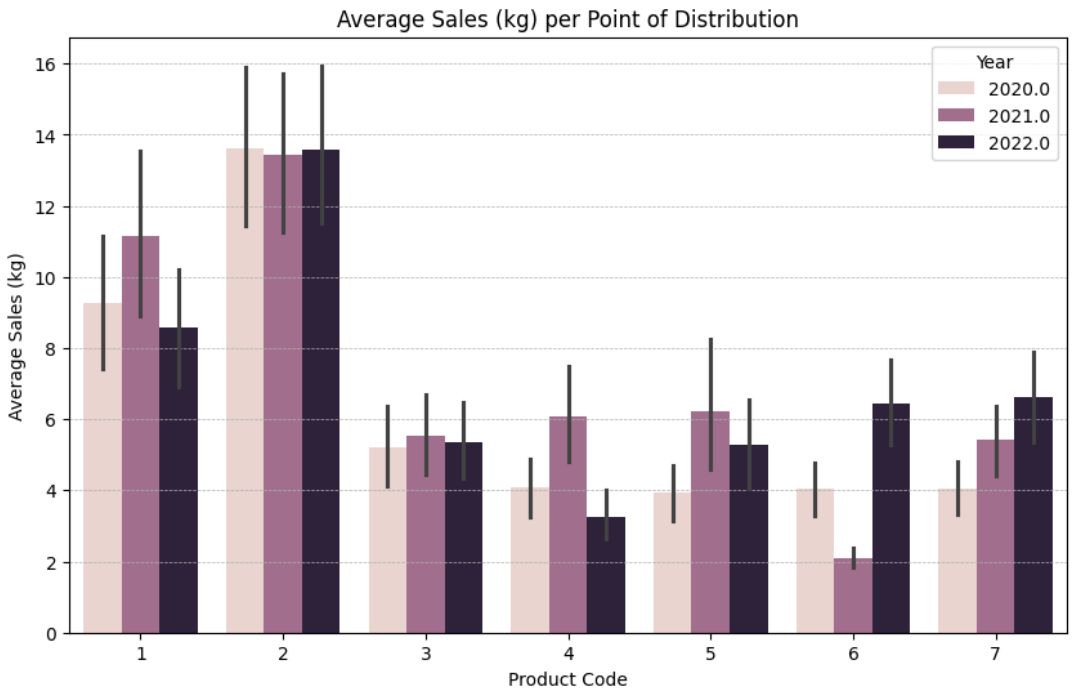

Figure 3. Average Sales per Point of Distribution

Figure 3 depicts the average sales per point of distribution. The x-axis shows the product code and y-axis shows the average sales in kilograms. The analysis includes three years such as 2020, 2021 and 2022. It shows 7 product codes. Where product 2 have high sales in three years when compare to other six products with range of 14.

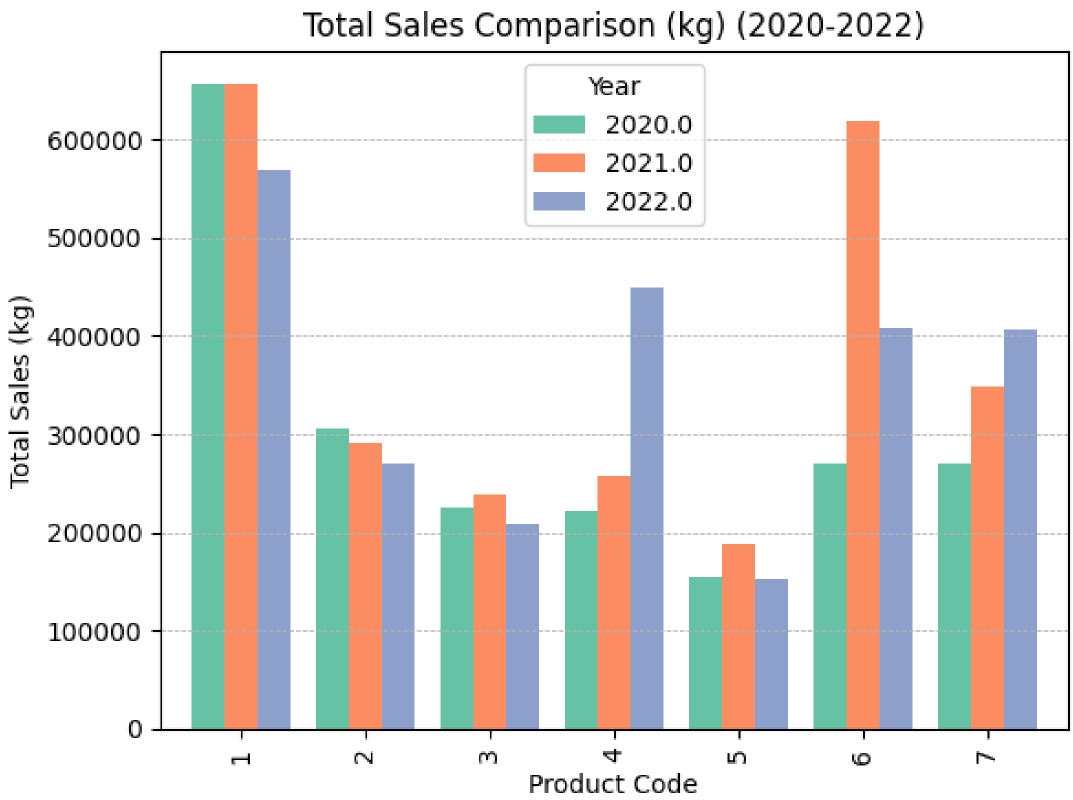

Figure 4. Comparison of Total Sales

Figure 4 depicts the comparison of total sales. The x-axis represents the product code from 1 to 7 and y-axis represents total sales in kilograms. The analysis is made for the year from 2020 to 2022. Where the product 1 reach peaks in total sales for all the three years when compare to other products and the sales is greater than 600000.

4.3 Performance Analysis

Table 1. Performance Analysis of MSE

Models	MSE
Random Forest Regressor	291630.46
Linear Regression	335009.24
Support Vector Regression	569211.36
Decision Tree Regression	562510.25
Proposed Model	**36011.6**

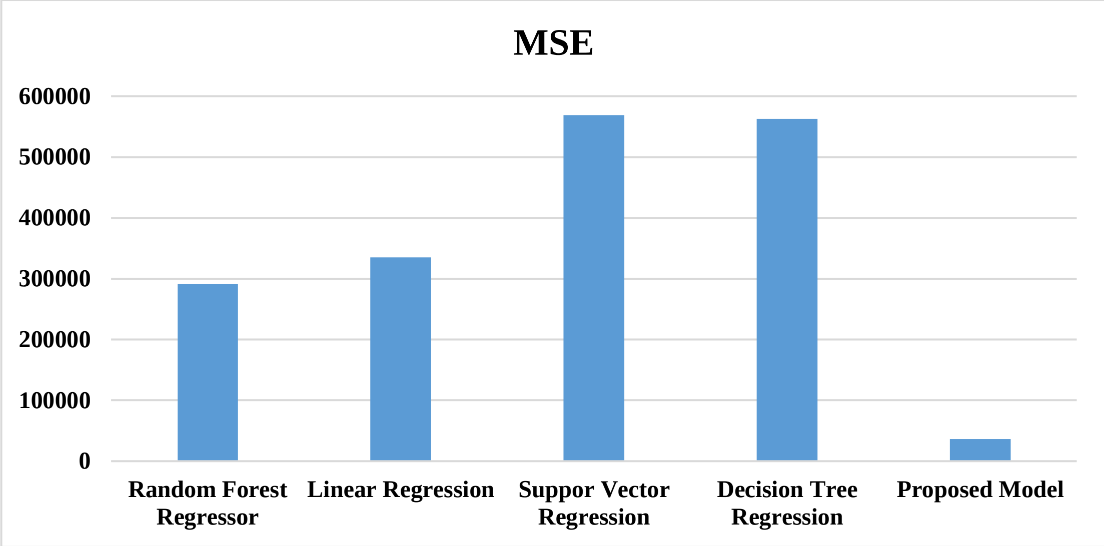

Figure 5. Performance Analysis of MSE

Table 1 represents the performance analysis of proposed model with other regression models such as random forest regressor, linear regression, support vector regression and decision tree regression in terms of MSE performance. The proposed model achieves better performance with 36011.6 when compare to other models with 291630.46, 335009.24, 569211.36 and 562510.25. It shows that the proposed model outperforms than other models. The graphical representation of table 1 is represented in figure 5.

Table 2. Performance Analysis of R^2

Models	R^2
Random Forest Regressor	0.8579
Linear Regression	0.8368
Support Vector Regression	0.7227
Decision Tree Regression	0.726
Proposed Model	**0.9824**

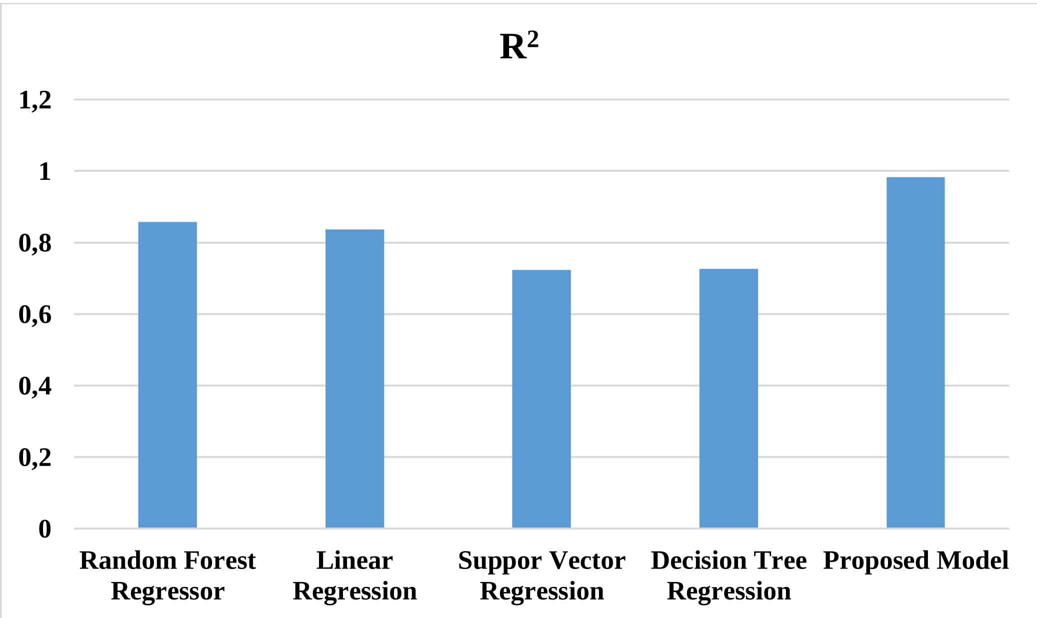

Figure 6. Performance Analysis of R^2

Table 2 represent the performance analysis of proposed model with other regression models such as random forest regressor, linear regression, support vector regression and decision tree regression in terms of R^2 performance. The proposed model achieves best performance with 0.9824 when compare to other models with 0.8579, 0.8368, 0.7227 and 0.726. It shows that the proposed model achieves ideal performance than other models. The graphical representation of table 2 is represented in figure 6.

Table 3. Performance Analysis of MAE and RMSE

Models	MAE	RMSE
Random Forest Regressor	252	540.02
Linear Regression	314.9	578.79
Support Vector Regression	278.19	754.46
Decision Tree Regression	427.56	750
Proposed Model	**82**	**189.76**

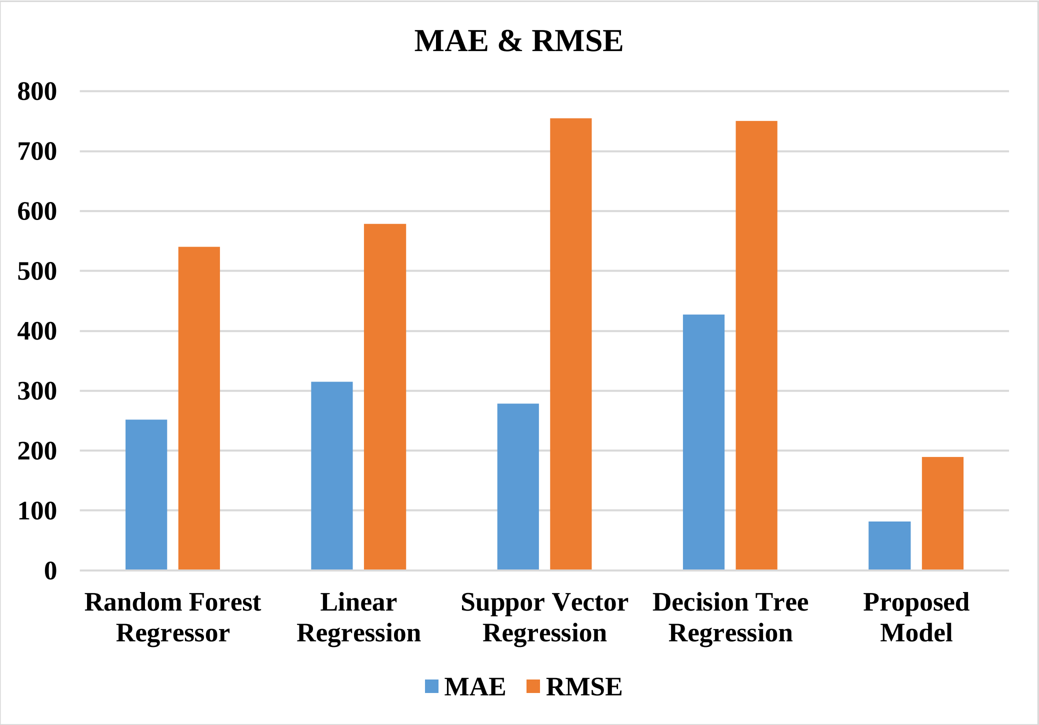

Figure 7. Performance Analysis of MAE and RMSE

Table 3 represents the performance analysis of proposed model with other regression models such as random forest regressor, linear regression, support vector regression and decision tree regression in terms of MAE and RMSE performance. The proposed model achieves best performance with 82 and 189.76 when compare to other models with the MAE values of 252, 314.9, 278.19 and 427.56. The RMSE value of other models are 540.02, 578.79, 754.46 and 750. It shows that the proposed model outperforms than other models. The graphical representation of table 1 is represented in figure 7.

5. Conclusion and Future Recommendation

In dairy industries, sales prediction is one of the major task. This might assist to neglect the returns loss and acquire about the high sale locations. Several prevailing studies have used different methods and techniques to predict the dairy sales, while those techniques have struggled with some technical issues. To tackle this problem, the proposed research has employed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model to forecast the dairy sales. The proposed research has been evaluated by the dairy supply chain sales dataset. The proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model has been compared with four regression models namely random forest regressor, linear regression, support vector regression and decision tree regression. The performance of the proposed Multi-Perspective Fusion Bi-LSTM with Universal Scale CNN model has evaluated by performance metrics of RMSE, MAE, MSE and R^2 with 189.76, 82, 36011.6 and 0.9824.

This represented that the proposed models has outperformed than other existing models with ideal performance. The future work of the proposed research will be forecast the sales in different companies from different countries of dairy industries

6. Declaration

- **Conflict of Interest:** The author reports that there is no conflict of Interest.

- **Funding:** None
- **Acknowledgement:** None

- **Data Availability Statement:** My manuscript has no associated data.

Reference

[1] R. Huerta-Soto, E. Ramirez-Asis, J. Tarazona-Jiménez, L. Nivin-Vargas, R. Norabuena-Figueroa, M. Guzman-Avalos*, et al.*, "Predictable inventory management within dairy supply chain operations," 2023.

[2] B. G. Hansen, Y. Li, R. Sun, and I. J. E. S. w. A. Schei, "Forecasting milk delivery to dairy–How modern statistical and machine learning methods can contribute," vol. 248, p. 123475, 2024.

[3] M. Azizsafaei, A. Hosseinian-Far, R. Khandan, D. Sarwar, and A. J. S. Daneshkhah, "Assessing risks in dairy supply chain systems: a system dynamics approach," vol. 10, p. 114, 2022.

[4] A. Goli, H. Khademi-Zare, R. Tavakkoli-Moghaddam, A. Sadeghieh, M. Sasanian, and R. J. N. c. i. n. s. Malekalipour Kordestanizadeh, "An integrated approach based on artificial intelligence and novel meta-heuristic algorithms to predict demand for dairy products: a case study," vol. 32, pp. 1-35, 2021.

[5] M. Bennett, "Statistical Forecasting for a Dairy Company: A Study to Determine a Best-Fit Model," Utica University, 2022.

[6] J. Sayyad, K. Attarde, B. J. B. D. Yilmaz, and C. Computing, "Improving Machine Learning Predictive Capacity for Supply Chain Optimization through Domain Adversarial Neural Networks," vol. 8, p. 81, 2024.

[7] Z. Li, A. Zuo, and C. J. S. Li, "Predicting Raw Milk Price Based on Depth Time Series Features for Consumer Behavior Analysis," vol. 15, p. 6647, 2023.

[8] B. Ji, T. Banhazi, C. J. Phillips, C. Wang, and B. J. b. e. Li, "A machine learning framework to predict the next month's daily milk yield, milk composition and milking frequency for cows in a robotic dairy farm," vol. 216, pp. 186-197, 2022.

[9] K. Karwasra, G. Soni, S. K. Mangla, Y. J. I. J. o. L. R. Kazancoglu, and Applications, "Assessing dairy supply chain vulnerability during the Covid-19 pandemic," vol. 27, pp. 2378-2396, 2024.

[10] A. Khanna, S. Jain, A. Burgio, V. Bolshev, and V. J. F. Panchenko, "Blockchain-enabled supply chain platform for Indian dairy industry: Safety and traceability," vol. 11, p. 2716, 2022.

[11] S. Abualuroug, A. Alzubi, and K. J. A. S. Iyiola, "Inventory Prediction Using a Modified Multi-Dimensional Collaborative Wrapped Bi-Directional Long Short-Term Memory Model," vol. 14, p. 5817, 2024.

[12] N. Slob, C. Catal, and A. J. P. v. m. Kassahun, "Application of machine learning to improve dairy farm management: A systematic literature review," vol. 187, p. 105237, 2021.

[13] S. Esmailizade, A. Ebrahimi, H. Soltani, A. Sam, and M. J. A. P. Rahimi, "Machine Learning Approaches for Retail Forecasting: A Study on XGBoost and Time-Series Models," 2024.

[14] S. Badruddoza, A. C. Carlson, and J. J. J. A. McCluskey, "Long-term dynamics of US organic milk, eggs, and yogurt premiums," vol. 38, pp. 45-72, 2022.

[15] M. J. A. Cockburn, "Application and prospective discussion of machine learning for the management of dairy farms," vol. 10, p. 1690, 2020.

[16] P. Chongstitvatana, "Demand Forecasting in Production Planning for Dairy Products Using Machine Learning and Statistical Method."

[17] G.-J. Streefland, F. Herrema, and M. J. S. A. T. Martini, "A Gradient Boosting model to predict the milk production," vol. 6, p. 100302, 2023.

[18] M. Alwadi, A. Alwadi, G. Chetty, and J. J. I. J. o. I. T. Alnaimi, "Smart dairy farming for predicting milk production yield based on deep machine learning," vol. 16, pp. 4181-4190, 2024.
[19] D. Piwczyński, B. Sitkowska, M. Kolenda, M. Brzozowski, J. Aerts, and P. M. J. A. S. J. Schork, "Forecasting the milk yield of cows on farms equipped with automatic milking system with the use of decision trees," vol. 91, p. e13414, 2020.
[20] S.-i. Shibata, Y. Kashiwazaki, T. Shimauchi, H. J. S. i. S. Kimura, and Technology, "Improving milk sales quantitative estimation by using POS data," vol. 9, pp. 61-69, 2020.
[21] M. Malik, V. K. Gahlawat, R. S. Mor, V. Dahiya, and M. J. L. Yadav, "Application of optimization techniques in the dairy supply chain: a systematic review," vol. 6, p. 74, 2022.
[22] F. Zhang, K. Weigel, and V. J. J. o. D. S. Cabrera, "Predicting daily milk yield for primiparous cows using data of within-herd relatives to capture genotype-by-environment interactions," vol. 105, pp. 6739-6748, 2022.
[23] H. Radwan, H. El Qaliouby, E. A. J. J. o. A. V. Elfadl, and A. Research, "Classification and prediction of milk yield level for Holstein Friesian cattle using parametric and non-parametric statistical classification models," vol. 7, p. 429, 2020.
[24] M. Bovo, M. Agrusti, S. Benni, D. Torreggiani, and P. Tassinari, "Random Forest Modelling of Milk Yield of Dairy Cows under Heat Stress Conditions," vol. 11, p. 1305, 2021.
[25] D. J. Innes, L. J. Pot, D. J. Seymour, J. France, J. Dijkstra, J. Doelman*, et al.*, "Fitting mathematical functions to extended lactation curves and forecasting late-lactation milk yields of dairy cows," vol. 107, pp. 342-358, 2024.
[26] K. S. Themistokleous, N. Sakellariou, E. J. C. Kiossis, and E. i. Agriculture, "A deep learning algorithm predicts milk yield and production stage of dairy cows utilizing ultrasound echotexture analysis of the mammary gland," vol. 198, p. 106992, 2022.
[27] E. Rana, A. K. Gupta, A. Singh, A. P. Ruhil, R. Malhotra, S. Yousuf*, et al.*, "Prediction of first lactation 305-day milk yield based on bimonthly test day milk yield records in Murrah buffaloes," 2021.
[28] P. R. J. C. Shorten and E. i. Agriculture, "Computer vision and weigh scale-based prediction of milk yield and udder traits for individual cows," vol. 188, p. 106364, 2021.
[29] M. Bovo, M. Agrusti, S. Benni, D. Torreggiani, and P. J. A. Tassinari, "Random forest modelling of milk yield of dairy cows under heat stress conditions," vol. 11, p. 1305, 2021.
[30] A. J. A. Atalan, "Forecasting drinking milk price based on economic, social, and environmental factors using machine learning algorithms," vol. 39, pp. 214-241, 2023.
[31] A. Liseune, M. Salamone, D. Van den Poel, B. Van Ranst, M. J. C. Hostens, and E. i. Agriculture, "Leveraging latent representations for milk yield prediction and interpolation using deep learning," vol. 175, p. 105600, 2020.
[32] C. G. Frasco, M. Radmacher, R. Lacroix, R. Cue, P. Valtchev, C. Robert*, et al.*, "Towards an Effective Decision-making System based on Cow Profitability using Deep Learning," in *ICAART (2)*, 2020, pp. 949-958.
[33] R. Huerta-Soto, E. Ramirez-Asis, J. Tarazona-Jiménez, L. Nivin-Vargas, R. Norabuena-Figueroa, M. Guzman-Avalos*, et al.*, "Predictable inventory management within dairy supply chain operations," *International Journal of Retail & Distribution Management,* vol. 53, pp. 1-17, 2025.
[34] A. Q. Olufemi-Phillips, O. C. Ofodile, A. S. Toromade, N. Abbey Ngochindo Igwe, and L. Eyo-Udo, "Utilizing predictive analytics to manage food supply and demand in adaptive supply chains," *Journal of Agricultural Economics and Management,(pending publication),* 2024.
[35] M. Alwadi, A. Alwadi, G. Chetty, and J. Alnaimi, "Smart dairy farming for predicting milk production yield based on deep machine learning," *International Journal of Information Technology,* vol. 16, pp. 4181-4190, 2024.
[36] M. Bovo, M. Agrusti, L. Ozella, C. Forte, D. Torreggiani, and P. Tassinari, "A viable data driven method for the assessment of the productivity level of dairy cows in future lactations," *Computers and Electronics in Agriculture,* vol. 230, p. 109860, 2025.

[37] D. T. Lianou, Y. Kiouvrekis, C. K. Michael, N. G. Vasileiou, I. Psomadakis, A. P. Politis*, et al.*, "The Use of Explainable Machine Learning for the Prediction of the Quality of Bulk-Tank Milk in Sheep and Goat Farms," *Foods,* vol. 13, p. 4015, 2024.

[38] V. Vasanthraj, V. Potdar, H. Agrawal, and A. Kaur, "Transforming milk supply chains with blockchain: Enhancing visibility and cost reduction," *Benchmarking: An International Journal,* vol. 32, pp. 1686-1719, 2025.

[39] D. Sharma, G. Singh, and A. Kaur, "Blockchain-based layered architecture for traceability in dairy supply chain management," *International Journal of Information Technology,* pp. 1-8, 2025.

[40] M. Malik, V. K. Gahlawat, R. S. Mor, and M. K. Singh, "Unlocking dairy traceability: Current trends, applications, and future opportunities," *Future Foods,* p. 100426, 2024.

[41] P. Vern, A. Panghal, R. S. Mor, V. Kumar, and S. Jagtap, "Blockchain-based traceability framework for agri-food supply chain: a proof-of-concept," *Operations Management Research,* pp. 1-20, 2024.

[42] D. Tran, F. Melissari, S. Le Feon, A. Papadakis, T. Zahariadis, D. Chatzitheodorou*, et al.*, "A holistic assessment of blockchain-based traceability systems: The case of protected designation of origin feta cheese production," *Computers and Electronics in Agriculture,* vol. 230, p. 109863, 2025.

[43] Y. Kazancoglu, M. D. Sezer, M. Ozbiltekin-Pala, Ç. Lafçı, and P. Sarma, "Evaluating resilience in food supply chains during COVID-19," *International Journal of Logistics Research and Applications,* vol. 27, pp. 688-704, 2024.

[44] N. Ali and W. Shah, "Predicting Retail Sales for Walmart: A Comprehensive Study Integrating Machine Learning and Time Series Models," 2024.

[45] Y. Bai, "Machine Learning Implementation for Demand Forecasting in Supply Chain Management," 2025.