Available at e-Journal Universitas Sultan Ageng Tirtayasa

JOURNAL INDUSTRIAL SERVICESS

journal homepage: http://jurnal.untirta.ac.id/index.php/jiss

eISSN 2461-0631

Research article

Cluster-specific Bi-LSTM models for improved pharmaceutical sales forecasting

Indri Hapsari

Department of Industrial Engineering, Universitas Surabaya, Jl. Raya Kalirungkut, Surabaya 60299, Jawa Timur, Indonesia

ARTICLE INFO

Article history:
Received 27 June 2024
Received in revised form 19 December 2024
Accepted 16 January 2025
Published online 15 January 2025

Keywords:
Forecasting
Bi-LSTM
K-means clustering
Pharmacy
Machine Learning

Editor:
Bobby Kurniawan

Publisher's note:
The publisher remains neutral concerning jurisdictional claims in published maps and institutional affiliations.

ABSTRACT

In the pharmaceutical sector, accurate forecasting is imperative due to the diverse range of medicines. Inadequate inventory levels risk patient well-being, whereas excessive stock can result in financial waste. This study categorizes historical sales data from a Denpasar-based pharmacy industry into clusters via K-Means clustering, analyzing 106 medicines over 60 months, with 54 months for training and six for testing. Four distinct Bi-LSTM (Bidirectional Long Short-Term Memory) based forecasting models emerged, tailored to each cluster's characteristics. Cluster 0's model, with two input neurons and three hidden layers, underwent 220 epochs of training, achieving an MAE of (0.0835) and an MSE of (0.0165). Cluster 1's model, more intricate with ten input neurons and two hidden layers, was trained for 136 epochs, resulting in an MAE of (0.1299) and an MSE of (0.0309). Cluster 2's model resembled Cluster 0 but with reduced neurons in the hidden layers, trained for 20 epochs, yielding an MAE of (0.0899) and an MSE of (0.0380). Finally, Cluster 3's model featured two input neurons and a single hidden layer with 128 neurons, trained for 150 epochs, attaining an MAE of (0.0239) and an MSE of (0.094). Forecast application of Cluster 0's model for individual medicine using Bi-LSTM demonstrated its efficacy in predicting demand compared with machine learning forecast models such as Random Forest, Gradient Boosting, Support Vector Machine, and Neural Network. The model's adaptability to demand fluctuations can guide pharmacies in managing their inventory and optimizing supply chain operations, sales, marketing strategies, and product development.

1. Introduction

The pharmaceutical industry has the most robust research and development and high supply chain costs. With the expiration of patents and the increase in the production of generic medicines, pharmaceutical companies must focus on developing more efficient and effective supply chains to face the challenges of demand forecasting and supply management [1]. Kochakkashani et al. [2] also discussed planning the supply chain for medicines and vaccines during the COVID-19 pandemic, which includes cold and non-cold chains and various types of medicines and vaccines. The optimization method creates a demand forecast for medicine in four stages, aiming to minimize costs in the medicine supply chain while maintaining service levels. The pharmaceutical industry is vital in providing health medicines and medicines the public needs. Rapidly increasing demand and the complexity of the factors influencing consumption patterns drive more careful planning and management of supplies. Accurate demand forecasting is essential to maintain adequate medicine availability for patients. In contrast, inaccuracies in forecasting can lead to shortages of stocks that can treat the patients' fatality and health, and excess stock can cause a loss for the company. Developing a reliable forecasting model is significant for maintaining the smooth production and distribution of medicine and optimizing inventory management.

The pharmacy industry's big data and machine learning have become increasingly important in the post-COVID-19 era and the industrial revolution 4.0. Big data generated by various sources, such as electronic medical records and patient data, provides deeper insight into public health trends; by combining machine learning capabilities, the accuracy and responsiveness of pharmaceutical forecasting to rapid changes in the healthcare environment increase. Actual practical application in pharmaceutical forecasting has proven that machine learning can identify patterns that are difficult to detect by conventional methods. This update enables healthcare providers and

*Corresponding author:
Email: indri@staff.ubaya.ac.id

http://dx.doi.org/10.62870/jiss.v10i2.26898

pharmaceutical companies to anticipate demand for medicine, manage supply chains more efficiently, and take better steps to respond to global health crises such as the COVID-19 pandemic. Machine learning for forecasting also allows personalization of care based on individual data, ensuring more precise and effective treatment [3].

This study uses the K-means clustering and Bi-LSTM (Bidirectional Long Short-Term Memory) methods as an approach that can improve the accuracy of demand forecasting in the pharmaceutical industry, especially during the Pharma 4.0 period. Pharma 4.0 is a concept that represents the industry's adoption of digitalization-enabled techniques and processes, cloud computing, the Internet of Things, and big data to gain a competitive advantage in domestic and global markets [4] . By grouping based on similar patterns, the K-means clustering method identifies groups of data with similar demand behavior, enabling a more focused focus on specific forecasts and thereby increasing the accuracy and effectiveness of demand forecasting. The K-Means clustering method serves as a data segregation approach that groups pharmaceutical sales data into distinct clusters based on historical demand patterns. By grouping data into clusters with similar characteristics, the K-Means method improves the accuracy of demand forecasting, allowing for the creation of a tailored forecasting model for each cluster. This method effectively captures the unique variations and characteristics of each product group, thereby enhancing the accuracy and efficiency of the predictive model. As highlighted by Kochakkashani et al. [2], clustering methods such as K-Means play a significant role in classifying different types of drugs and vaccines in the supply chain. This is because the K-Means method has a structured framework to handle data heterogeneity. Ouedraogo [6] talks about how K-Means can help choose better machine learning models by effectively clustering, and Palupi and Fakhruzzaman [7] show how it can be used to divide pharmacies into groups so that marketing strategies work better. The flexibility and computational efficiency of the K-Means method make it very useful in the pharmaceutical industry, as it reveals patterns and trends needed in demand forecasting and decision-making. This study applies the K-Means clustering method to classify drugs into clusters based on historical sales patterns, which enables the development of a Bi-LSTM model specifically tailored to each cluster's characteristics. This approach provides a solid foundation for achieving accurate and responsive forecasting results. On the other hand, Bi-LSTM is one of the most effective neural network architectures for modeling time series data. Its main advantage is the ability to capture long-term dependencies in data, responding to challenges that often arise in demand forecasting, such as high fluctuations and irregular patterns. Pharmaceutical sales data uses clustering to address its inherent complexity and diversity. We group medicines based on similar demand patterns. This clustering allows for the creation of an appropriate demand forecasting model, thereby improving the accuracy and efficiency of demand prediction, as well as simplifying data characteristics. This approach also improves inventory management and resource allocation. The application of clustering in this study bridges the gap between data heterogeneity and the need for insight, thus providing a solid foundation for further forecasting using the Bi-LSTM model. The combination of K-means clustering and Bi-LSTM methods makes the approach superior in increasing demand forecasting accuracy. This combination will provide optimal results, especially when enough data or big data can be processed.

This paper's contribution positively impacts the development of more sophisticated and accurate forecasting methods and has the potential to become an essential reference for professionals in the pharmaceutical industry facing increasingly complex and dynamic demand forecasting challenges. Using trend-based prediction results, companies can analyze factors influencing demand fluctuations and identify exploitable opportunities—ethical medicine requiring a doctor's prescription falls under this category. Most people with specific diseases usually need several medicine variants to help heal because the symptoms experienced are more than one symptom. Based on this, the pattern of selling medicine to each other has the potential to influence each other. The ability of Bi-LSTM to predict trends of increasing and decreasing demand for a type of medicine variant can be an indicator for companies to analyze other medicine variants that are likely to be affected by the trend.

In the pharmaceutical industry, data clustering has the potential to reveal patterns or correlations arising from complex data, such as clinical trial results, patient data, product data, and the like. Applying clustering data in the pharmaceutical industry brings significant benefits, especially in classifying patients based on their disease profile or response to treatment. By grouping patients based on similar characteristics, pharmaceutical companies can identify groups of patients who are likely to respond similarly to certain medicines or therapies. This approach can advance the development of more focused and personalized treatment for each patient. In addition, data clustering also plays a role in the development and discovery of new medicines. Researchers can identify compounds with similar characteristics by grouping clinical trials and molecular data from various medicine compounds. This identification can assist in selecting potentially more efficient and effective medicine candidates. However, data clustering in the pharmaceutical industry also faces several challenges. Choosing the correct clustering method, interpreting accurate results, and handling large and complex data requires sophisticated technological infrastructure and in-depth analysis skills.

In conclusion, this paper proposes a novel approach for predicting demand in the pharmaceutical industry, leveraging the combined power of K-Means clustering and Bi-LSTM models. The pharmaceutical sector faces unique challenges in demand forecasting due to the diverse range of medicines and the critical importance of maintaining optimal inventory levels. The study systematically analyzes historical sales data from a Denpasar-based pharmacy industry, categorizing medicines into clusters using K-Means clustering. Subsequently, four distinct Bi-LSTM-based forecasting models are developed, each tailored to the specific characteristics of its respective cluster.

The contributions of this research are threefold:

1. It addresses the pressing need for accurate demand forecasting in the pharmaceutical industry, emphasizing the potential consequences of inadequate inventory levels on patient well-being and financial waste.
2. The paper introduces an innovative methodology that combines data clustering and deep learning to enhance forecasting precision. This dual approach allows for a more focused, personalized forecasting strategy based on identified clusters.
3. The study provides detailed insights into the performance of each forecasting model, highlighting their adaptability to diverse demand patterns within the pharmaceutical sector.

The remainder of the paper is structured as follows. Section 2 provides an overview of the pharmaceutical industry's challenges, emphasizing the need for sophisticated demand forecasting methods. This section also outlines the methodology, detailing the steps taken in K-Means clustering and the development of the Bi-LSTM model. Section 3 presents the experiment's results, showcasing the proposed approach's effectiveness through metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE). This section also compares the Bi-LSTM models with conventional machine learning forecast models, demonstrating their superior ability to capture intricate temporal dependencies. Finally, in Section 4, the paper summarizes the key findings and discusses the broader implications of the proposed approach for the pharmaceutical industry.

2. Material and method

Big data, which refers to large, complex, and diverse data sets, provides significant benefits in this industry. Researchers can utilize clinical trials, patient records, and epidemiological data to comprehend diseases, disease patterns, and patient responses to treatment. In addition, big data also supports the development of new medicines through the analysis of genetic, biological, and molecular data to identify potential therapeutic targets and improve treatment effectiveness. Big data also provides advantages in increasing the efficiency of the medicine supply chain and distribution. Pharmaceutical companies can optimize stocks, reduce logistics costs, and identify demand patterns by analyzing logistics and inventory data. Ramesh and Santhi [5] describe the exploration of big data analytics in the healthcare industry, highlighting the potential of big data analytics and machine learning in addressing future healthcare challenges and grouping big data to form clusters with similar characteristics. Clustering is grouping documents, observations, or cases into classes with certain similarities. Each class or cluster contains data that is like each other and different from other clusters. The main goal of clustering is understanding and use. Understanding includes the initial stage of grouping, which can be followed by summarization (mean, standard deviation) and assigning a class label to each group for supervision-based classification purposes. The purpose of use focuses more on finding prototypes representing each group representatively, thus providing an abstraction of each data object.

Previous studies have addressed demand forecasting in the pharmaceutical industry using machine learning or time series models without incorporating clustering techniques. Some studies try to guess how much a drug will sell by using simple Recurrent Neural Networks (RNNs) or Long Short-Term Memory (LSTM) networks. These networks look at the whole dataset and don't consider the differences between parts of the dataset, like drugs with high versus low demand variability [2, 6, 7]. Some researchers use alternative clustering techniques, such as DBSCAN or hierarchical clustering, but these methods are often computationally intensive or poorly suited for segmenting large-scale sales data based on demand patterns [9], [10]. In addition, many forecasting models adopt a generic approach, applying the same network architecture to the entire dataset, resulting in suboptimal performance [13], [14].

This research model integrates K-Means clustering and Bi-LSTM to improve the accuracy of pharmaceutical sales forecasting. The process begins with historical sales data. We then apply the K-Means clustering method to group the data into distinct clusters based on sales patterns, with each cluster representing drugs that share similar demand characteristics. By handling each cluster independently, this clustering enables a more focused and tailored forecasting approach. Next, the data for each cluster is fed into a Bi-LSTM model that is specifically designed to account for its unique patterns and temporal dependencies. The Bi-LSTM model leverages its ability to handle sequential data, capturing the complex relationships between past and future sales trends. Finally, the model generates precise sales forecasts for each cluster, which collectively contribute to better and more efficient inventory management in pharmaceutical supply chain operations. This model is strong enough to handle the challenges of changing and varied demand in the pharmaceutical industry because it combines the segmentation power of K-Means with the predictive power of Bi-LSTM.

2.1. K-means clustering

K-means clustering, a versatile and widely employed unsupervised learning algorithm, finds notable applications in the pharmaceutical industry. It is a pivotal tool for categorizing pharmaceutical products based on historical demand patterns, enabling tailored approaches for optimized forecasting and decision-making. Applying K-means clustering spans diverse domains, demonstrating its versatility and effectiveness. Ouedraogo [6] employs K-means to enhance machine learning model selection for interpretable breast cancer diagnosis, achieving high performance and interpretability. Palupi and Fakhruzzaman [7] utilizes a hybrid algorithm combining swarm intelligence and K-means to segment pharmacy retailers, informing tailored marketing strategies. Murry [8] employs K-means clustering to identify patient-centered care preferences, revealing distinct clusters with preferences for autonomy and collaboration in pharmacist care. Setiawan et al. [9] optimizes hospital clustering in Jakarta post-COVID-19 using K-means, concluding its appropriateness for the dataset and highlighting its dependence on dataset characteristics. These studies underscore K-means clustering's adaptability and effectiveness in addressing various challenges across healthcare, marketing, and decision-making contexts. In the pharmaceutical field, Mousavi et al. [10] use principal component analysis and clustering analysis to understand the life cycle of pharmaceutical companies in Iran. Kochakkashani et al. [2] apply the K-means clustering algorithm to classify different types of medicines and vaccines in the supply chain, while Arief et al. [11] combine the clustering approach with the Pharma 4.0 concept to understand digital transformation in the Indonesian pharmaceutical industry. For the conditions of the COVID-19 pandemic, Devaraj et al. [12] and Luo et al. [13] used a time series method based on deep learning to predict the number of COVID-19 cases worldwide and analyze the development of the pandemic. Another method from Ilu and Prasad [14] used integrated statistical significance techniques and K-means clustering to improve the ARIMA prediction model in predicting COVID-19 cases. Time series and K-means clustering are techniques used to group similar data points. Time series clustering is specifically tailored to data organized in a temporal sequence, offering a specialized approach compared to the general applicability of K-means clustering. Adapting K-means for time series data may not capture temporal dependencies as effectively as dedicated time series clustering algorithms designed for sequential data. K-means clustering offers several advantages for forecasting medical demand. Firstly, it efficiently partitions the data into distinct clusters based on demand patterns, allowing for a more granular understanding of consumer behavior. This method facilitates the identification of similar demand trends within clusters, aiding in predicting future demand for specific medicine. Additionally, K-means clustering is computationally efficient, enabling rapid analysis and adaptation to evolving demand scenarios. This approach enhances demand forecasting accuracy, contributing to more effective inventory management and resource allocation in the pharmaceutical industry.

The Elbow Method is commonly used for cluster analysis to determine the optimal number of clusters. The main focus of this method is to look for points on the graph that show significant changes in the explanation of data variation when the number of clusters changes. The Elbow Method aims to find the number of clusters that can explain the data variation well without being too complex or simple. If the number of clusters is too small, the data may not be appropriately classified, while if the number of clusters is too large, the clustering results can lose meaning or be challenging to interpret. Therefore, the Elbow Method helps find the correct number of clusters for analysis. Another method, the Silhouette Method, measures how well each data point fits into a predetermined cluster and evaluates the distinctiveness of the cluster from others. The essence of the Silhouette Method lies in identifying the number of clusters that yield effective clustering results and maintain a clear separation between clusters without significant overlap. This method considers the clustering quality for each data point, not just the overall variation in the data.

2.2. *Neural network*

Research by Hole et al. [15] states that the digitization process in the pharmaceutical industry includes the increased use of robotics, automation, and computerization to reduce costs, increase efficiency and productivity, and be flexible to change. Other researchers, such as Arden et al. [16], research the application of Industry 4.0 in pharmaceutical manufacturing and preparation for future smart factories. The application of these technologies has the potential to drastically increase the speed, efficiency, flexibility, and quality of medicine production. This research discusses the evolution of pharmaceutical manufacturing from manual processes with simple tools to large pharmaceutical industries with advanced technology. Kim et al. [17] face the challenge of developing a new digital value chain model with a product-life cycle approach in the biopharmaceutical industry. Advances in Pharma 4.0 technology in biopharmaceuticals complicate and slow the development of new medicines. This study proposes integrating the pharmaceutical value chain model with a product life cycle perspective to understand changes in medicine development and the application of digital transformation at different stages of medicine development. In the Saudi Arabian pharmaceutical manufacturing sector, Halwani et al. [18] explored nanotechnology as an opportunity for pharmaceutical companies to compete in the growing global medicine market by developing innovative medicines, especially nanotechnology. Another researcher, Duarte et al. [19], created decision-support tools for equitable and sustainable medicine distribution. Arief et al. [11] apply the Pharma 4.0 concept to the pharmaceutical industry in Indonesia. This research identifies eight core competencies needed by human resources in the pharmaceutical industry to succeed in this era: critical thinking, digital skills, and data ethics. Regarding implementing digital Pharma 4.0, this study identified five main levels ranging from simplification to disease prediction.

To prevent the spread of COVID-19 through the development of medicines and vaccines and the study of epidemic diffusion, Lee and Chang [20] proposed a Stochastic Susceptible-Infected-Quarantine-Removed (SIQR) model with two delays. The research complemented its theoretical findings with numerical simulations using the Gillespie algorithm. The study results indicate that the proposed SIQR model better explains the spread of the COVID-19 epidemic compared to the traditional Susceptible-Exposed-Infected-Released (SEIR) model regarding MAPE, RMSE, and MAD. This article contributes significantly to the mathematical understanding of infectious disease spread and is valuable for health scientists and epidemiological researchers in planning disease control strategies. Furthermore, as heart disease prediction becomes increasingly important due to a rising number of sudden deaths caused by heart disease, Chandrasekar [21] employs artificial neural networks, specifically the Convolutional Neural Network (CNN) VGG-19 model. This research leverages machine learning, deep learning, and data mining technologies. VGG-19 CNN and Deep Neural Network (DNN) architecture successfully classify histopathological images for heart disease prediction. Hoffmann and Rutschmann [22] specifically link big data analysis with demand forecasting in the supply chain. This research shows that companies can use big data analysis in operational, tactical, or strategic demand planning, provide better results in forecasting accuracy, and consider many parameters directly affecting customer demand. By incorporating Big Data analytics, companies can significantly enhance forecasting accuracy by taking into account numerous parameters directly influencing customer demand. This comprehensive approach allows for a more nuanced understanding and effective response to the dynamic nature of market demands at different planning levels.

Artificial Neural Network (ANN) is used in forecasting to increase accuracy. This method can overcome nonlinear relationships between various parameters, not only limited to statistical and numerical data in the past. ANN is an information processing system that resembles a biological neural network developed as a mathematical model of human cognition or neurobiology. ANN represents system dynamics as a mathematical model, determining the accuracy of network architecture, network learning techniques, and network models. Understanding system behavior patterns allows the application of these results to predict future behavior. One of the ANN variants specifically designed to deal with sequential or time series data is the Recurrent Neural Network (RNN). RNNs can understand temporal relationships in data because of the repeated connections that allow information to be stored and passed between time steps. However, traditional RNNs have constraints such as vanishing gradient and burst problems, which limit their ability to capture long-term dependencies. To address this, more sophisticated RNN variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have emerged, incorporating specific control mechanisms to manage the flow of information over time.

2.3. Bi-LSTM

LSTM is a model that can consider past information over time, enabling it to recognize patterns, trends, and dependencies in time series data. These advantages make it a good choice for data modeling complex time frame structures. With a "cell state" as an internal memory unit, the LSTM can store and access long-term information. Control gates control how much information is remembered or forgotten in the forecast. The strengths of LSTM in forecasting include the ability to handle remote dependencies in data, overcome the vanishing gradient problem common to traditional RNNs, and provide more accurate predictions for complex time sequences. LSTM forecasts various time series variables, including product sales, financial data, and energy use. With its ability to model complex time frame structures, LSTM provides more reliable predictions and supports decision-making regarding demand planning and management. However, applying LSTM in forecasting requires understanding model configuration, proper data processing, and validating the results obtained. In practice, practitioners must adapt the selection of forecasting models and methods to the data context and specific forecasting objectives. Research conducted by Lin et al. [23] and Shubo et al. [24] discusses the use of deep learning models based on LSTM and Mutual Information (MI) for short-term electricity load predictions.

Bidirectional LSTM (Bi-LSTM) is a Recurrent Neural Network (RNN) type that effectively understands and predicts patterns in time series data. The advantage of the Bi-LSTM method is its ability to "look" backward and forwards from a data sequence to better accommodate the previous context and approach forecasts. The application of Bi-LSTM has been successful in various studies related to forecasting. Zhou et al. [25] developed a COVID-19 prediction model using an LSTM-based deep-learning approach. The application of the LSTM was also successfully carried out in the research by Khalil et al. [26], where a combined model of Convolutional Neural Networks and LSTM is used to forecast public transport demand. The Bi-LSTM method has also provided promising results in other studies. Rayan and Alaerjan [27] applied Bi-LSTM in processing X-ray images to identify COVID-19 infection. Sirisha et al. [28] used the Deep Stacked Bi-LSTM model to predict oil production. Arief et al. [11] used the Deep Bi-LSTM model to predict multivariate time series data features using sensors. Unlike the traditional Feed-Forward Neural Network, the Bi-LSTM Neural Network has internal nodes at each layer that are not connected. In the hidden layer connection, a directional loop allows the recall and storage of previous information in memory units, thereby enhancing the association between data segments in different time series. The neural network combines the previous and current input to determine its output. Another advantage of the Bi-LSTM Neural Network is its ability to consider the forward and backward correlations of a given time series data that will improve its performance.

Bi-LSTM has two main components, namely forward LSTM and backward LSTM. In the forward section, the system processes data from the past to the future, while in the backward section, it processes data from the future to the past. Integrating information from both parts of the LSTM enhances the accuracy and reliability of predictions. Each LSTM has internal memory cells, which allow them to store and manipulate information over a more expansive period. LSTM cells can add, remove, or update information based on new input and knowledge drawn from the past. Bi-LSTM is very effective in the case of sequential or time series data, such as in language prediction, classification, and modeling. The results from the forward and backward LSTMs are combined to make predictions on the next time step. Utilization of time series data and its feedback provides additional context that helps the model understand the problem better and faster. Pustohkhina et al. [29] used a hyperparameter search convolutional neural network with Bi-LSTM (HPS-CBL) to use big data to detect intrusion. Experimental results suggest that the HPS-CBL model performs better than the compared methods, with maximum levels of precision and accuracy. Using the Information Processing View theoretical framework, Ziaee et al. [30] use Big Data Analytics in the Pharmaceutical Supply Chain in Australia. Metlek [31] developed and tested a new deep learning model to predict aircraft fuel consumption using Convolutional Neural Network (CNN) and Bi-LSTM. The study compares the developed hybrid model with existing deep learning models, specifically LSTM and Bi-LSTM. Kota and Munisamy [32] use Neural Network-based Deep Learning for sentiment analysis by combining Convolutional Neural Network, Bi-LSTM, and attention mechanisms. CNN helps reduce complexity, and Bi-LSTM helps process long input text sequences.

2.4. Method

This study uses the K-means clustering method as the first stage to classify medicine sales data based on historical sales patterns. The research procedure began with collecting sales data from a pharmacy company in Denpasar for 17 medicines in antibiotics, antifungals, and antivirals for 60 months from January 2017 to December 2021. This data includes sales information per month for each medicine. After the data is collected, the second step involves preprocessing the data to remove missing values or significant values that may affect the clustering results. The third step is identifying and resolving missing or invalid values in the data. Finally, if necessary, outlier detection and handling are carried out. This step will eliminate potential interference with forecasting results. The following process is to apply the K-means clustering method to group these medicines based on their historical sales patterns. At this stage, leverage clustering methods like k-means, Hierarchical clustering, or DBSCAN. The primary objective is to identify groups of medicines exhibiting similar behavior. In K-means clustering, managing the number of clusters is crucial, as this parameter determines how the data will be partitioned into distinct groups. The number of clusters must be determined based on the nature of the data and the complexity of the problems encountered.

Classifying products into specific groups, the team forms a demand forecasting model for each product group utilizing the Bi-LSTM method. The subsequent step entails training these models with 54 months of sales data for each product group. This training process is tailored to optimize the models' performance, aiming for accurate forecasts. To ensure accuracy, the forecasting model then undergoes evaluation and validation using test data, which comprises six months of sales data. When using Bi-LSTM, parameters such as the number of layers, neurons, and the activation function must be considered. The number of layers in an LSTM network controls the complexity of the model. However, adding fewer layers can result in overfitting. The number of neurons in each layer also affects the model's ability to capture patterns in the data. This model is updated through a training process with historical data to optimize its performance. The activation function in each LSTM layer regulates how information flows in the network. A structured experimental design and accurate parameter determination are required to compare the K-means clustering and Bi-LSTM methods in forecasting medicine demand. To evaluate performance, we test both methods using test data and assess forecasting accuracy by comparing the results with actual data. If deviations are present, we analyze factors affecting forecasting performance and implement corrective measures, such as model adjustments or parameter changes. In this comparison, we can refer to evaluation metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE).

Finally, we implemented this model to ensure the maintenance of forecast quality over time. The model is updated regularly according to changes in data and operational conditions. This approach ensures that forecasting remains accurate and relevant in dealing with changes in the pharmaceutical industry. Thus, integrating K-means clustering and Bi-LSTM methods provides a solid foundation for better and more effective demand forecasting in this sector. The presentation of demand forecasting results will display a visualization of medicine demand forecasting. Time series charts allow the identification of trends, seasonal patterns, and fluctuations in data from time to time. At the end, we present conclusions and recommendations based on forecasting results. These recommendations cover determining which method is more appropriate to apply in the context of a pharmaceutical company and the rationale behind it. The data sources and references related to the implemented K-means clustering and Bi-LSTM methods need to be mentioned to maintain the integrity of the information. Figure 1 illustrates the K-Means and Bi-LSTM research model. The process begins with data preprocessing, followed by K-Means clustering, which segments the time series data into clusters based on similar patterns. The clustering results are then used to configure the training process, which includes initialization, parameter setup, and layer adjustment for the Bi-LSTM model. The model undergoes iteration and testing to minimize errors, with further parameter optimization if needed. Once it achieved error minimization, the model was moved to the testing configuration and terminated it based on the number of epochs. We validate the resulting forecasts using MAE and MSE metrics and finally produce the visualization of prediction results to illustrate the model's performance.

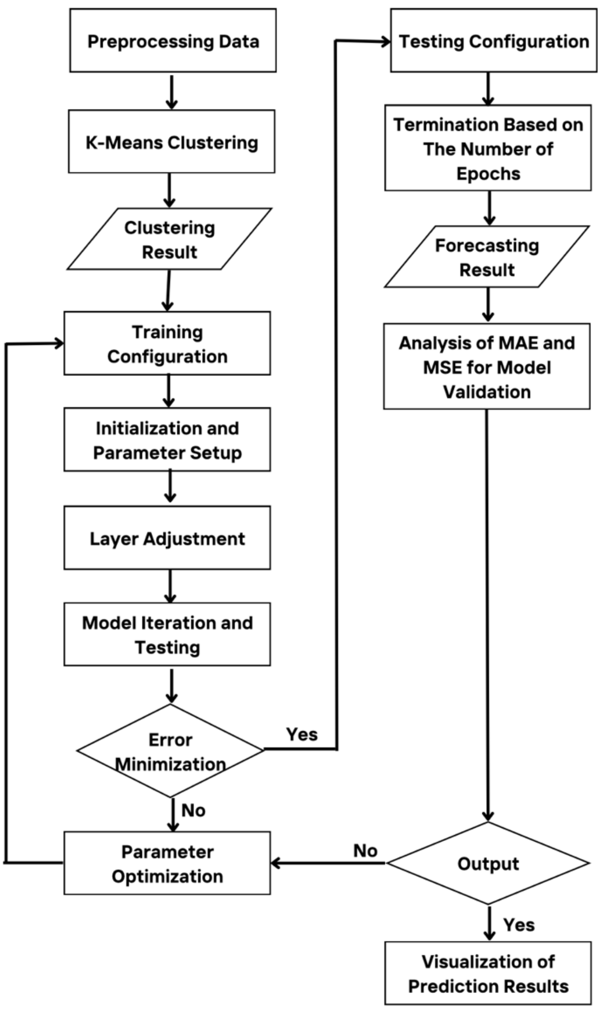

Figure 1. K-Means and Bi-LSTM Model

3. Results and discussions

We will use medicine sales data for analysis, recorded on a monthly time series basis. This data represents the total sales for each medicine, excluding the number of returns. The data used is medicine sales data for the category of antibiotics, antifungals, and antivirals for 60 months, with data sharing in the form of 54 months of training data and six months of testing data. The total number of medicines marketed varies year to year because several times there were discontinues for medicines with low sales levels and the addition of new medicines. In the first year, there were 162 medicines; the following year, 179, 181, 178 and 161 medicines. During this preprocessing stage, we reduced the total number of medicines marketed in the first year to 106. The medicines selected to be the object of research are for the categories of antibiotics, antifungals, and antivirals, which have historical data of 60 months. This range was chosen to obtain sales data patterns before the COVID-19 pandemic, during the COVID-19 pandemic, and during the new standard transition period. The forecasting results aim to predict the current sales pattern of each medicine, considering that medicines encompass items whose sales were affected by the pandemic. The medicine code consists of 5 digits, with the first two digits representing information on the dosage form of the medicine, for example, in the form of an injection, tablet, syrup, and so on, and the other three digits being the code for each medicine.

The method used in this study is time K-means clustering with Python. After obtaining results from the clusters and forming several groups, we will create an appropriate model to predict the members of each cluster. In the data preprocessing stage, we reduced the number of medicine variants in the antibiotic, antifungal, and antiviral categories 16 by excluding medicines without complete historical data for 60 months, from January 2017 to December 2021. The data processing for K-means clustering utilized training data, consisting of sales transaction data for 54 months from 16 medicine variants in the antibiotics, antifungal, and antiviral categories, which will serve as input for the forecasting process. The first stage in conducting K-means clustering analysis is to perform data overviewing and then determine the optimal number of clusters with the Elbow Method in Figure 2, Silhouette Method in Figure 3, and Gap Statistics in Figure 4.

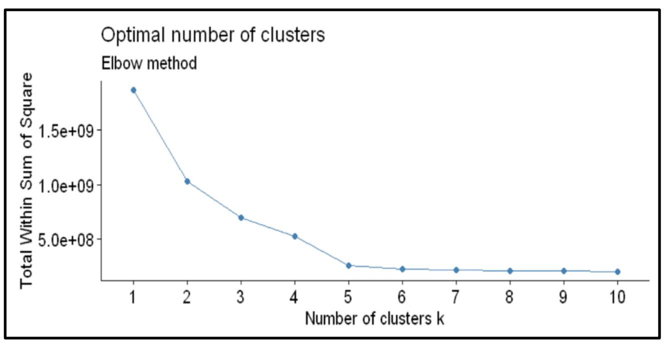

Figure 2. Elbow Method

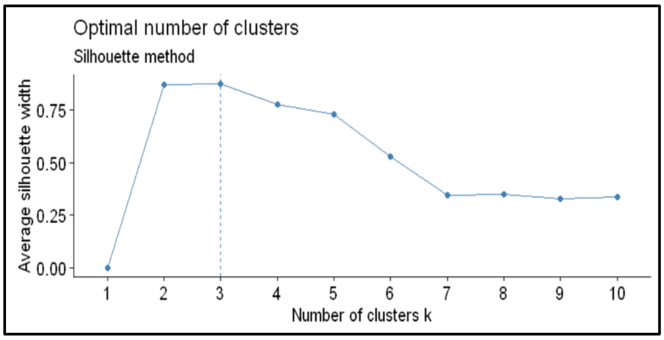

Figure 3. Silhouette Method

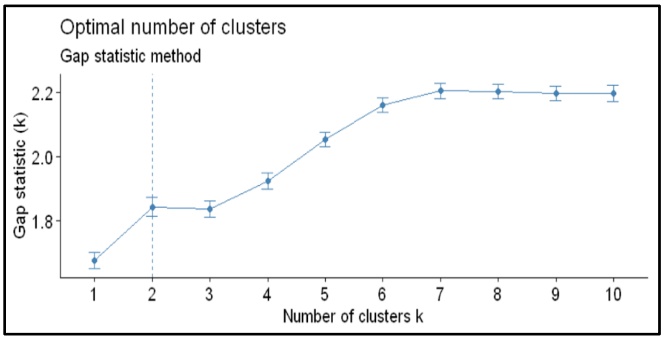

Figure 4. Gap Statistical Method

In the Elbow Method, the optimum number of clusters k suitable for this research dataset is 3. In the Silhouette Method, the optimal number of clusters is k. The recommended number is 3 clusters. The optimum number of clusters k produced in the Gap Statistics method is 2. However, unfortunately, the optimum number of cluster k produced by each method is different and does not provide a definite output, so we decided to use an additional method, NBClust. NBClust is a library that contains 30 algorithms for determining the number of clusters. cluster-k to make it easier for users to determine the correct number of cluster-k. The reason for adding this method is that NBClust is a library that can summarize all the results of determining the number of cluster-k from 30 algorithms. The NBClust Method analysis reveals that most algorithms, which aim to determine the optimal number of clusters denoted as 'k,' converge on the conclusion that the ideal number of clusters is 3. We grouped the medicines into three clusters with the distribution of cluster members: 101 medicines in the first cluster, four medicines in the second cluster, and 1 medicine in the third cluster.

Cluster 1 has too many members and needs to match with other clusters. Fahim et al. (2008) assert in their journal that the quality of clusters diminishes when the dataset exhibits a spherical shape with notable variations in size. So, we will split Cluster 1 back into Cluster 0 and Cluster 1. The medicines in cluster 1 have an average monthly sale of 116.79 medicines. Many medicines have monthly sales figures lower than 116.79. Consequently, we will further divide the cluster into cluster 0, including medicines with an average monthly sales figure of less than 116.79, and the remaining medicines will form cluster 1. The medicine distribution shows that 60.8% of medicines fall into Cluster 0, 28.5% fall into Cluster 1, 0.6% fall into Cluster 2, 1.9% fall into Cluster 3, and 8.4% fall into Cluster 4.

The second stage is to forecast the data using an Artificial Neural Network algorithm of the Bi-LSTM type with Python. The train data used is medicine transaction sales data for 54 months since January 2017. The test data covered the last six months, from July 2021 to December 2021. A forecasting system model uses the average data for each medicine in the cluster. The respective Bi-LSTM models for each cluster and the output from the resulting data processing are as follows:

In cluster 0, the Bi-LSTM model used is Bi-LSTM with an input layer with two neurons, the first hidden layer with 32 neurons, the second with 32 neurons, and the third with 16 neurons with one output layer. The number of epochs (training) used was 220 times. After completing the training process, Figure 5 and Figure 6 display visualizations of the MAE and MSE from the model. The validation loss and training loss curves, derived from the MAE and MSE accuracy methods, converge at the minimum error value, ensuring the model's suitability for the applied data. Figure 7 provides a visualization of the forecast outcomes for cluster 0, marking the completion of the forecasting system model for this cluster.

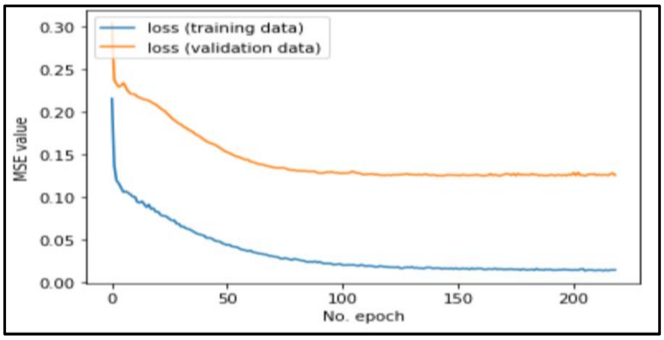

Figure 5. Cluster 0 MSE plot

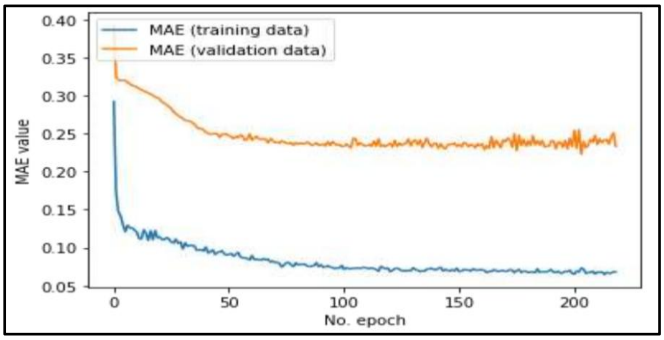

Figure 6. MAE Cluster 0 plot

In cluster 1, the Bi-LSTM model features input layers with ten neurons, a first hidden layer consisting of 128 neurons, a second hidden layer with 32 neurons, and an output layer. The training process involves 136 epochs. For cluster 2, the Bi-LSTM model is characterized by an input layer comprising two neurons, the first hidden layer with 32 neurons, the second with eight neurons, the third with four neurons, and an output layer, with training conducted over 20 epochs. In cluster 3, the Bi-LSTM model utilizes an input layer with two neurons, a hidden layer of 128 neurons, and an output layer with one neuron, undergoing training for 150 epochs. All the forecast validation results showed that every cluster's MAE and MSE values range from 0 to ∞ with recommendations for good MAE and MSE values close to 0. Cluster Analysis 0 reveals four distinct clusters based on sales levels. In Cluster 0, comprising 70 medicines with the lowest sales, the average monthly sales stand at 41,645 medicines—the Bi-LSTM data processing results in an MAE of 0.0835 and MSE of 0.0165 during the forecasting validation. Moving to Cluster 1, which includes 31 medicines with sales surpassing Cluster 0 but remaining relatively low at an average of 286.5 medicines per month, the Bi-LSTM forecasting validation yields an MAE of 0.1299 and MSE of 0.0309. Cluster 2, representing four medicines with the highest sales across clusters, boasts an average monthly sales figure of 2027 medicines, resulting in a Bi-LSTM forecasting validation with an MAE of 0.0899 and MSE of 0.0380. Lastly, Cluster 3, housing a single medicine with a moderately high sales level akin to Cluster 2 (averaging 806.2 medicines monthly), reports a Bi-LSTM forecasting validation with an MAE of 0.0239 and MSE of 0.0094.

The advantage of the Bi-LSTM method is that it helps understand sequential patterns in time series data well. This ability allows the model to capture complex relationships between past and future data, which is relevant in predicting the demand for medicine that tends to follow specific sequential patterns. This method can also handle time series data with varying sequential lengths to allow for more flexible and efficient forecasting without the rigid standardization of time lengths. Bi-LSTM can also handle multivariable data input well. Forecasting demand for medicine allows the model to consider interactions between various factors affecting demand, such as seasonality, promotions, and special events. The disadvantage is that this deep-learning model requires a lot of data to train correctly. If medicine demand data is limited or sparse, this model may not provide optimal results. In addition, training and evaluating the Bi-LSTM model requires high computational processing, especially if the time series data is extensive. This extension can lead to long training times and requires powerful computational resources. Bi-LSTM also has high complexity, which can lead to the risk of overfitting the training data if not managed properly. Overfitting can cause the model to be unusual and unreliable in forecasting new data.

The forecasting landscape has witnessed significant advancements in integrating machine learning models, offering enhanced accuracy and efficiency. This paper explores the strengths of various machine learning models in forecasting and provides a concise overview of four prominent models: Random Forest, Gradient Boosting, Support Vector Machines (SVM), and Neural Networks. Random Forest is an ensemble learning technique that constructs numerous decision trees during training and averages their predictions, making it robust for large datasets with high dimensionality and minimizing overfitting. Gradient Boosting, another ensemble method, builds trees sequentially to address residual errors, excels in capturing complex data relationships, demonstrates resilience to outliers, and offers high predictive accuracy. Support Vector Machines (SVM) serve as powerful models for classification and regression tasks, operating by mapping data into a higher-dimensional space and finding optimal hyperplanes to maximize class margin. SVMs are effective in high-dimensional spaces and versatile across various data types. Neural Networks, especially deep learning models, have gained popularity for automatically learning complex patterns. Comprising interconnected layers of artificial neurons, Neural Networks excel in capturing intricate relationships, particularly in tasks with substantial data, making them effective for time series forecasting.

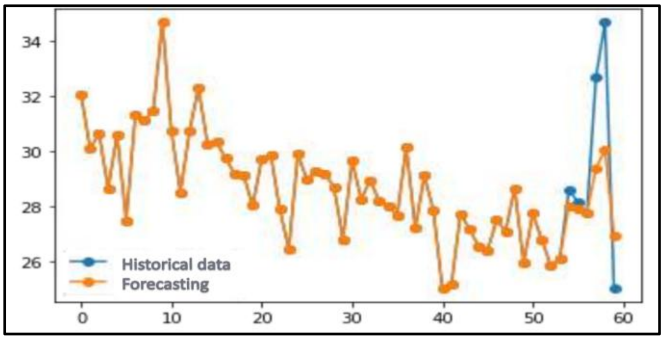

Figure 7. Visualization of Cluster 0 Forecasting Results

While these models have demonstrated remarkable forecasting capabilities, this paper also delves into the unique features of Bi-LSTM networks. Bi-LSTM, a recurrent neural network (RNN), handles sequential and time-dependent data by maintaining long-term dependencies. It introduces bidirectionality, allowing the model to capture information from both past and future timestamps, enhancing its ability to understand complex temporal patterns. In comparing these models, the paper explores their commonalities and differences, emphasizing how Bi-LSTM, with its specialized architecture, addresses challenges inherent in time series forecasting. Through a comprehensive evaluation, the paper aims to provide insights into the optimal selection of machine learning models based on specific forecasting requirements and dataset characteristics.

We applied the forecasting model to each medicine in cluster 0. One of the medicines that is a member of cluster 0 is a medicine with code 04040. Figure 8 is a forecasting plot produced using the forecasting model formed for cluster 0. The Bi-LSTM models employed for forecasting in different clusters exhibit specific configurations.

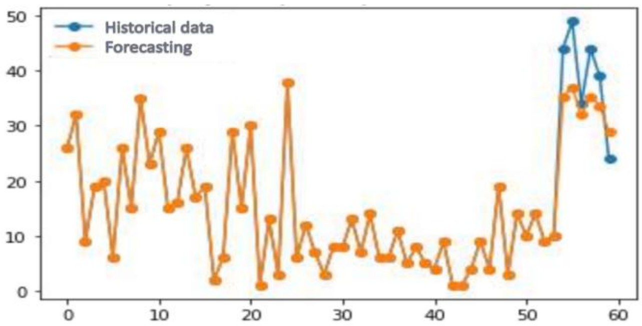

Figure 8. Visualization of Medicine 04040 Forecasting Results

The comparison of Mean Squared Error (MSE) and Mean Absolute Error (MAE) across different forecasting models reveals distinct performance metrics for each method for medicine 04040. Among the models evaluated using Python, Bi-LSTM stands out with remarkably superior results compared to other techniques. Regarding MSE, Bi-LSTM achieved a significantly lower value of 2.733583, indicating its ability to minimize the squared differences between predicted and actual values. This result outperformed other models such as Random Forest (5.291067), Gradient Boosting (8.523067), SVM (9.223233), and Neural Networks (68.677367). The lower MSE for Bi-LSTM suggests a higher accuracy in capturing the variability within the data and generating more precise forecasts. Similarly, when considering MAE, which represents the absolute differences between predicted and actual values, Bi-LSTM demonstrated exceptional accuracy with a value of 1.495000. This MAE value is notably lower than the MAE values of other models, including Random Forest (2.080000), Gradient Boosting (2.613333), SVM (2.513333), and Neural Networks (8.060000). The smaller MAE for Bi-LSTM underscores its effectiveness in providing more reliable point forecasts.

The superior forecasting performance of Bi-LSTM can be attributed to its unique architecture, allowing it to capture complex temporal dependencies and patterns within the data. The model's ability to remember and learn from sequential information, combined with its bidirectional processing, contributes to enhanced predictive capabilities. While other models may struggle to capture intricate relationships, Bi-LSTM excels in handling sequential data, making it a favorable choice for time-series forecasting tasks. The observed differences in MSE and MAE highlight the importance of choosing appropriate models based on the dataset's specific characteristics and the forecasting problem's nature.

4. Conclusions

In this study, four Bi-LSTM models were developed for each cluster, each with distinct architectures tailored to the characteristics of the corresponding dataset. The resulting MAE and MSE visualizations indicated successful model training, with validation and training loss curves reaching their lowest points. The forecasting system model for Cluster 0 was then applied to individual medicines within the cluster, such as the medicinal with code 04040. The results show the strengths of Bi-LSTM among various machine learning models in forecasting of four models: Random Forest, Gradient Boosting, Support Vector Machines (SVM), and Neural Networks.

Pharmaceutical companies can use K-means clustering and Bi-LSTM methods to improve demand forecasting accuracy for their medicine. Higher forecasting accuracy will help companies plan and manage inventory more efficiently, avoid shortages or excess inventories, and increase customer satisfaction. Various factors, including market trends, seasons, marketing campaigns, and other external factors, often influence the demand for medicine in the pharmaceutical industry. Companies can be more responsive to changes in demand and anticipate demand fluctuations that may occur. Pharmaceutical companies can optimize medicine production and distribution processes with more accurate forecasting. This optimization will assist in more efficient production arrangements, avoid wastage, and reduce operational costs.

To enhance the quality of medicine demand forecasting, researchers can explore several potential future topics for further development. These topics include paying attention to the characteristics of medicine, such as generic and innovative medicines. Additionally, research can focus on understanding customer behavior through customer behavior analysis and developing predictive analysis techniques to identify factors influencing purchasing decisions, considering aspects like promotion, season, and medical developments. It is necessary to combine forecasting methods such as K-means clustering Bi-LSTM, and other methods such as AutoRegressive Integrated Moving Average (ARIMA) and Exponential Smoothing to provide better results. The continued development of Artificial Intelligence (AI) and machine learning technologies opens up new opportunities for developing more sophisticated forecasting models. Deep learning and reinforcement learning techniques can improve a model's ability to identify complex patterns and predict demand more accurately.

Declaration statement

Indri Hapsari: **Writing-Original draft, Writing-Review & Editing**. Amelia Santoso: **Conceptualization, Methodology**. Ni Putu Pradnya Widyasari: **Writing-Original draft, Collecting and Processing data**.

Acknowledgement

The acknowledgement was given to anonymous refrees for constructive feedback.

Disclosure statement

The author declares that this manuscript is free from conflict of interest and is processed by applicable journal provisions and policies to avoid deviation from publication ethics in various forms.

Funding statement

The authors received no funding for this research.

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article.

AI Usage Statement

This manuscript utilizes generative AI and AI-assisted tools to improve readability and language. All AI-generated content has been reviewed and edited by the authors to ensure accuracy and scientific integrity. The authors take full responsibility for the content and conclusions of this work and disclose the use of AI to maintain transparency and comply with publisher guidelines.

References

[1] G. Merkuryeva, A. Valberga, and A. Smirnov, "Demand forecasting in pharmaceutical supply chains: A case study," *Procedia Comput. Sci.*, vol. 149, pp. 3-10, 2019, doi: 10.1016/j.procs.2019.01.100.

[2] F. Kochakkashani, V. Kayvanfar, and A. Haji, "Supply chain planning of vaccine and pharmaceutical clusters under uncertainty: The case of COVID-19," *Socio-Econ. Plan. Sci.*, vol. 87, 2023, doi: 10.1016/j.seps.2023.101602.

[3] E. Hofmann, H. Sternberg, H. Chen, A. Pflaum, and G. Prockl, "Supply chain management and industry 4.0: Conducting research in the digital age," *Int. J. Phys. Distrib. Logist. Manag.*, vol. 49, pp. 945-955, 2019, doi: 10.1108/IJPDLM-11-2019-399.

[4] I. C. Reinhardt, J. C. Oliveira, and D. T. Ring, "Current perspectives on the development of industry 4.0 in the pharmaceutical sector," *J. Ind. Inf. Integr.*, vol. 18, 2020, doi: 10.1016/j.jii.2020.100131.

[5] T. Ramesh and V. Santhi, "Exploring big data analytics in health care," *Int. J. Intell. Netw.*, vol. 1, pp. 135-140, 2020, doi: 10.1016/j.ijin.2020.11.003.

[6] D. N. Ouedraogo, "Interpretable machine learning model selection for breast cancer diagnosis based on K-means clustering," *Appl. Med. Inform. Original Res.*, vol. 43, no. 3, pp. 91-102, 2021.

[7] G. S. Palupi and M. N. A. Fakhruzzaman, "Indonesian pharmacy retailer segmentation using recency frequency monetary-location model and ant K-means algorithm," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 6, pp. 6132-6139, 2022, doi: 10.11591/ijece.v12i6.

[8] L. T. Murry, J. Urmie, and A. Olson, "Item development for patient-centered care preferences using a three-archetype heuristic: An exploratory study," *Explor. Res. Clin. Soc. Pharm.*, vol. 9, 2023, doi: 10.1016/j.rcsop.2023.100251.

[9] K. E. Setiawan, A. Kurniawan, A. Chowanda, and D. Suhartono, "Clustering models for hospitals in Jakarta using Fuzzy C-means and K-means," *Procedia Comput. Sci.*, vol. 216, pp. 356-363, 2023, doi: 10.1016/j.procs.2022.12.146.

[10] A. Mousavi, M. Mohammadzadeh, and H. Zare, "A clustering approach to identify the organizational life cycle," *J. Open Innov. Technol. Market Complex*, vol. 8, no. 3, 2022, doi: 10.3390/joitmc8030108.

[11] N. N. Arief, A. Gustomo, M. R. Roestan, A. N. A. Putri, and M. Islamiaty, "Pharma 4.0: analysis on core competence and digital leveling implementation in the pharmaceutical industry in Indonesia," *Heliyon*, vol. 8, no. 8, 2022, doi: 10.1016/j.heliyon.2022.e10347.

[12] J. Devaraj et al., "Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant?," *Results Phys.*, vol. 21, 2021, doi: 10.1016/j.rinp.2021.103817.

[13] Z. Luo, L. Zhang, N. Liu, and Y. Wu, "Time series clustering of COVID-19 pandemic-related data," *Data Sci. Manag.*, vol. 6, no. 2, pp. 79-87, 2023, doi: 10.1016/j.dsm.2023.03.003.

[14] S. Y. Ilu and R. Prasad, "Improved autoregressive integrated moving average model for COVID-19 prediction by using statistical significance and clustering techniques," *Heliyon*, vol. 9, no. 2, 2023, doi: 10.1016/j.heliyon.2023.e13483.

[15] G. Hole, A. S. Hole, and I. McFalone-Shaw, "Digitalization in the pharmaceutical industry: What to focus on under the digital implementation process?," *Int. J. Pharm.*, vol. 3, 2021, doi: 10.1016/j.ijpx.2021.100095.

[16] N. S. Arden, A. C. Fisher, K. Tyner, L. X. Yu, S. L. Lee, and M. Kopcha, "Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future," *Int. J. Pharm.*, vol. 602, 2021, doi: 10.1016/j.ijpharm.2021.120554.

[17] Y. Kim, E. Atukeren, and Y. Lee, "A new digital value chain model with PLC in biopharmaceutical industry: The implication for open innovation," *J. Open Innov. Technol. Market Complex*, vol. 8, no. 2, 2022, doi: 10.3390/joitmc8020063.

[18] A. A. Halwani et al., "Current status and vision of local pharmaceutical industries in Saudi Arabia: The focus on nano medicines," *Saudi Pharm. J.*, vol. 31, no. 8, 2023, doi: 10.1016/j.jsps.2023.06.007.

[19] I. Duarte, B. Mota, T. Pinto-Varela, and A. P. Barbosa-Póvoa, "Pharmaceutical industry supply chains: How to sustainably improve access to vaccines?," *Chem. Eng. Res. Des.*, vol. 182, pp. 324-341, 2022, doi: 10.1016/j.cherd.2022.04.001.

[20] H. S. Lee and B. Y. Chang, "A new stochastic SIQR (Susceptible-Infected -Quarantine-Removed) model with two delays: Forecasting COVID-19 diffusion," *Int.*

J. Ind. Eng.: Theory, Appl. Pract., vol. 30, no. 3, 2023, doi: 10.23055/ijietap.2023.30.3.8365.

[21] S. Chandrasekar, K. Subburathinam, and K. Srihari, “Analysis for heart disease prediction using deep neural network and VGG_19 convolutional neural network”, *Int. J. Ind. Eng.: Theory, Appl. Pract.*, pp. 876-889, 2023, doi: 10.23055/ijietap.2023.30.4.8603.

[22] E. Hofmann and E. Rutschmann, “Big data analytics and demand forecasting in supply chains: a conceptual analysis,” *Int. J. Logist. Manag.*, vol. 29, no. 2, pp. 739-766, 2018, doi: 10.1108/IJLM-04-2017-0088.

[23] J. Lin, J. Ma, J. Zhu, and Y. Cui, “Short-term load forecasting based on LSTM networks considering attention mechanisms,” *Int. J. Electr. Power Energy Syst.*, vol. 137, 2022, doi: 10.1016/j.ijepes.2021.107818.

[24] H. Shubo et al., “Short-Term load forecasting based on mutual information and BI-LSTM considering fluctuation in importance values of features,” *IEEE Access*, 2023, doi: 10.1109/ACCESS.2023.3323403.

[25] L. Zhou, C. Zhao, N. Liu, X. Yao, and Z. Cheng, “Improved LSTM-based deep learning model for COVID-19 prediction using optimized approach,” *Eng. Appl. Artif. Intell.*, vol. 122, 2023, doi: 10.1016/j.engappai.2023.106157.

[26] S. Khalil, C. Amrit, T. Koch, and E. Dugundji, “Forecasting public transport ridership: management of information systems using CNN and LSTM architectures,” *Procedia Comput. Sci.*, vol. 184, pp. 283-290, 2021, doi: 10.1016/j.procs.2021.03.037.

[27] A. Rayan and A. S. Alaerjan, “An improved crow search optimization with Bi-LSTM model for identification and classification of COVID-19 infection from chest X-Ray images,” *Alexandria Eng. J.*, vol. 76, pp. 787-798, 2023, doi: 10.1016/j.aej.2023.06.052.

[28] B. Sirisha, K. K. C. Goud, and B. T. V. S. Rohit, “A Deep Stacked Bidirectional LSTM (SBiLSTM) model for petroleum production forecasting,” *Procedia Comput. Sci.*, vol. 218, pp. 2767-2775, 2023, doi: 10.1016/j.procs.2023.01.248.

[29] I. V. Pustokhina, D. A. Pustokhin, and E. L. Lydia, “Hyperparameter search based convolution neural network with Bi-LSTM model for intrusion detection system in multimedia big data environment,” *Multimed. Tools Appl.*, vol. 81, pp. 34951-34968, 2022, doi: 10.1007/s11042-021-11271-7.

[30] M. Ziaee, H. K. Shee, and A. Sohal, “Big data analytics in the Australian pharmaceutical supply chain,” *Ind. Manag. Data Syst.*, vol. 123, no. 5, pp. 1310-1335, 2023, doi: 10.1108/IMDS-05-2022-0309.

[31] S. Metlek, “A new proposal for the prediction of an aircraft engine fuel consumption: a novel CNN-BiLSTM deep neural network model,” *Aircraft Eng. Aerosp. Technol.*, vol. 95, no. 5, pp. 838-848, 2023, doi: 10.1108/AEAT-05-2022-0132.

[32] V. R. Kota and S. D. Munisamy, “High accuracy offering attention mechanisms based deep learning approach using CNN/Bi-LSTM for sentiment analysis,” *Int. J. Intell. Comput. Cybern.*, vol. 15, no. 1, pp. 61-74, 2022, doi: 10.1108/IJICC-06-2021-0109.