A Survey on Clinical Time Series Forecasting Methods

Simi Job
University of Southern Queensland, Australia

Abstract—**Clinical time series forecasting is gaining research interest in recent times owing to its applicability in clinical decision support and personalised patient treatment. While conventional time series approaches such as ARIMA are regularly used for processing temporal data, deep learning methods are increasingly being adopted due to its ability in handling non linear data. Advances in the usage of temporal information has also led to the expansion of research in multivariate time series analysis. This work identifies the various methodologies adopted in time series analysis and the future development potential for researching temporal data.**

Index Terms—**ime series Multivariate Recurrent Neural Networks Convolutional Neural Networks.ime series Multivariate Recurrent Neural Networks Convolutional Neural Networks.T**

I. Introduction

The application of artificial intelligence systems in personalised healthcare is increasingly being researched to develop enhanced clinical decision support systems. The integration of various clinical data obtained from medical records, patient sensor devices, lab and diagnostic reports provide a pathway for developing a predictive and diagnostic system for personalised patient treatment. Efficiently harnessing the volumes of data available from these resources impart a well developed AI-enabled system for clinicians to act as a supplementary system for supporting personalised medicine.

Temporal data is used in various fields of study including Finance (D. Cheng, Yang, Xiang, & Liu, 2022), Power usage (Gasparin, Lukovic, & Alippi, 2019), Health (Yee, Narain, Akmaev, & Vemulapalli, 2019), (Staffini, Svensson, Chung, & Svensson, 2021), (Albers et al., 2018) and Climate (Afrifa-Yamoah, Mueller, Taylor, & Fisher, 2020). The information obtained from this data is useful for time series analysis, particularly in forecasting or classification tasks. Time series analysis of clinical data involves risk prediction of disease, forecasting of physiological values, etc, on the basis of past clinical values. Analysing temporal data provides insights into underlying trends and patterns in data over specific time periods.

Intensive care units are sources of high volumes of patient monitoring data. Utilisation of this information is minimal since it is physically impossible for healthcare workers to monitor and manually analyse variations of multiple physiologic variables. The substantial advances in artificial intelligence pave the way to develop intelligent systems that can assimilate ICU monitoring data along with lab data and other significant clinical values to provide forward insights for early interventional measures as well as forecasting resource demands. The availability of ICU data as temporal observations enable forecasting of physiological values based on historic data. These future values can then be assimilated to determine indicators of organ dysfunction or deterioration in terms of respiratory failure, cardiac distress, liver failure or progress towards septic shock. Several studies have utilised temporal ICU data for predicting sepsis shock (Yee et al., 2019), (Thao, Tra, Son, & Wada, 2018), cardiac stress (Yoon et al., 2019), survival (Thao et al., 2018), forecasting various physiological measures such as heart rate (Staffini et al., 2021), (Oyeleye, Chen, Titarenko, & Antoniou, 2022), blood pressure (E. Huang, Wang, Chandrasekaran, & Yu, 2020), blood glucose levels (Albers et al., 2018) or forecasting multiple physiologic values (Hamidi, Borzu, Maroufizadeh, & Amini, 2021).

The review is consolidated after researching relevant journals in the past five years. Articles before the year 2017 are excluded from this review, and therefore detailed information on conventional forecasting methodologies is not provided. Around three hundred articles were studied and the primary journals of search includes IEEE, Elsevier, ACM etc. From these research articles, 100 plus articles related to the topic were selected and reviewed.

This work seeks to study the most relevant models that are recently being used for time series modelling that are appropriate for the clinical domain. In addition to determining the recent approaches to time series analysis, the work also proposes to identify the various challenges associated with clinical time series forecasting. The review is expected to provide researchers with a direction towards methodologies adopted for clinical temporal analysis. The paper is organised as follows: Section II outlines the definition of clinical time series. Section III discusses the various methodologies in time series forecasting. Section IV presents the various challenges and issues that are encountered in time series forecasting. Section V discusses the trends and potential research directions. The survey is summarized in Section VI with Conclusion.

II. Time Series Analysis of Clinical Data

Researchers have used ICU data for time series classification (Karim, Majumdar, Darabi, & Harford, 2019) or forecasting (Staffini et al., 2021), (Oyeleye et al., 2022), (E. Huang et al., 2020), (Albers et al., 2018), (Hamidi et al., 2021). For instance, temporal analysis of various physiologic parameters such as blood pressure and lab data can be used to classify whether a patient is diabetic or not. Similarly, this data is also useful for forecasting values for adopting interventional measures.

Most of these studies are based on time series analysis with a univariate approach (Staffini et al., 2021), (Oyeleye et al., 2022), (E. Huang et al., 2020), (Albers et al., 2018), with some of the recent research focusing on multivariate time series (Hamidi et al., 2021). In multivariate time series analysis, a value is predicted based on the values of multiple variables. For this purpose, it is essential to identify correlation and dependencies among these variables.

Fig. 1 shows a representation of the process involved in forecasting multiple physiological values with reference to ICU data. Multiple variables or multiple time series are used for forecasting the variables. Training data is extracted from a set of data for an initial time period. Feature engineering is performed on this data and passed through the neural networks for learning the time series patterns. This learned information is tested on a latter part of time series data followed by forecasting for a future predefined period of time.

1) Problem Definitions: Time Series is a sequential set of observations measured on time points which may or may not be uniformly spaced. Univariate time series is the simplest form of time series and considers only one time-dependent variable. Multivariate time series analysis involves investigating inter-dependencies between multiple variables in a time series or between multiple related time series. It is considered to be more powerful in prediction due to the incorporation of relationship between different variables in the time series. Let $y_t = (y_{1t}, y_{2t}, \ldots y_{nt})$ be an n dimensional ICU time series. If y_{1t} and y_{2t} are the heart rate and blood pressure of a patient recorded at time point t respectively, the temporal dependence between these two variables can be analysed for the purpose of predicting the future heart rate value of the patient. The different variables in the time series represent patient information such as physiological, lab data or interventional measures captured at a specific time point. With a focus on accommodating to the interactivity between variables in a multivariate scenario, discretization of time series is often adopted (J. M. Lee & Hauskrecht, 2021). This involves converting the numbers in a time series into a group of discrete elements.

The focus of this review is on multivariate time series, wherein multiple physiological parameters such as blood pressure, heart rate etc. may be forecast based on past values. Multivariate time series analysis involves detecting periodic trends, identifying dependencies between variables or multiple time series, and anomaly detection. For two variables y_1 and y_2, forecasting for time t can be performed based on past n values. Computation of $y_1(t), y_2(t), y_n(t)$ considers the historic values of both y_1 and y_2. This can be represented as shown in Eqn. 1 and Eqn. 2. C_1 and C_2 are constants. z_1 is a coefficient and e_1, e_2 represent error factors.

$$y_1(t) = C_1 + z_1 11 * y_1(t-1) + z_1 12 * y_2(t-1) + e_1(t-1) \quad (1)$$

$$y_2(t) = C_2 + z_1 21 * y_1(t-1) + z_1 22 * y_2(t-1) + e_2(t-1) \quad (2)$$

2) Motivation: The primary motive for ICU temporal analysis is to establish a personalised treatment process for the patient. A large section of research in clinical time series forecasting has focused on univariate analysis. However, to obtain a reliable forecasting outcome it is imperative to incorporate multiple variables into the time series analysis process. As seen in Fig. 2, patient data includes ICU monitoring data, lab test results, diagnostic reports, medications administered and so on. While it is possible to forecast blood pressure values based only on historic blood pressure data, for developing a definitive personalised treatment, it is essential to include other data such as heart rate, glucose levels etc. (Seid et al., 2019) conducted a study to determine neonatal mortality risks with both univariate and multivariate approaches. While the univariate approach revealed that sepsis and hypothermia were not leading causes for mortality, multivariate analysis indicated these two factors were mortality predictors. Fig. 2 depicts the relevance of ICU and clinical data time series forecasting in personalised medicine. Significant information from ICU monitoring data and patient medical records are utilised for forecasting physiological values or probable lab results. These predicted values are availed to provide personalised treatment for a patient, such as adopting early intervention measures.

The review encompasses the univariate and multivariate approaches adopted in clinical temporal analysis and utilises the analytical results for developing a prediction system. The predicted values are used methodically by clinicians to assist in reliable decision making aided by AI enabled systems.

III. Methodologies in Time Series Forecasting

Time series forecasting has been conventionally performed with univariate statistical models such as ARMA, ARIMA or seasonal variations of these models. For multivariate time series, ARIMAX, VARMA etc. are used. However, statistical models are incapable of adapting to non-linearity. As a solution, deep learning methods are being increasingly used for multivariate time series analyses owing to their adeptness in handling non-linearity. This section discusses the various time series pre-processing methods and linear and non-linear methods used in time series analysis.

1) Time Series Decomposition: Time series data consists of trends, levels, seasonality factors and noise. Deconstructing the information presented in temporal data into several components aid in segregating the several patterns associated with each of these factors. The decomposition of time series as an additive model can be represented as Eqn. 3 where T_t, C_t, S_t and N_t represent the trend, cyclicity, seasonality and noise factors respectively.

$$y_t = T_t + C_t + S_t + N_t \quad (3)$$

The classical approach for time series decomposition is the Moving Average method in which the trend is computed by averaging the values in a specified time period. (Abdollahi, 2020) used empirical mode decomposition to build a hybrid model for price forecasting. The components were extracted as volatile and nonlinear classes. Markov-based model applied on the volatile components and SVM based forecasting on the nonlinear components were observed to attain the best performance results. The LSTM encoder based model

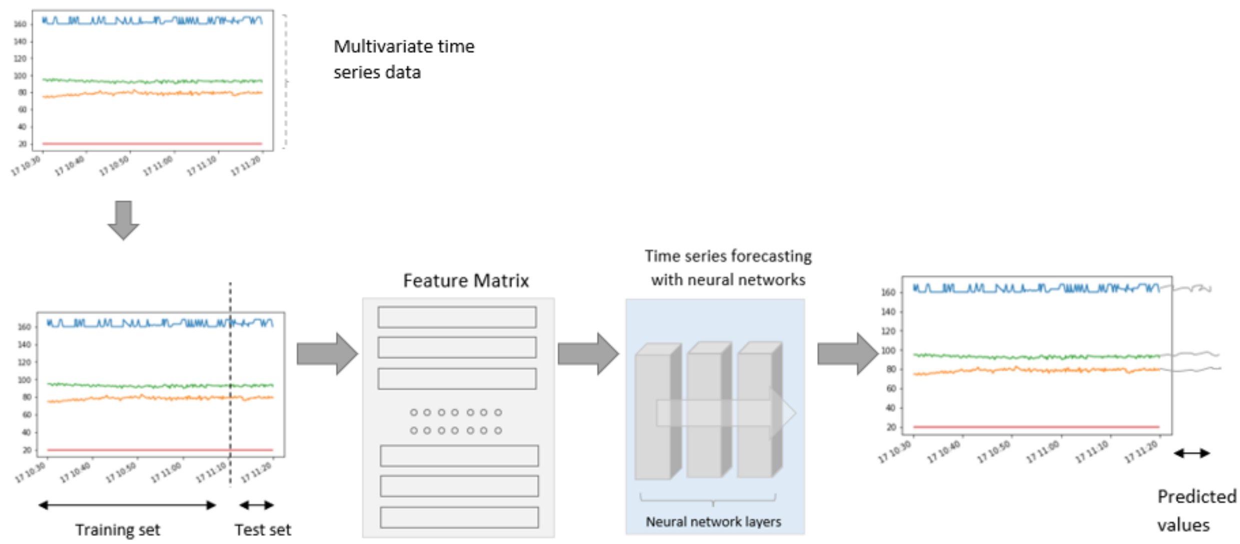

Fig. 1: Physiological value Prediction task pipeline

proposed by (Bedi & Toshniwal, 2020) incorporates variational mode decomposition technique for detecting significant features in conjunction with error variance modelling for improved performance. Decomposition is a methodology for enabling component wise analysis of temporal data and is used in various studies for highlighting the underlying temporal patterns.

2) Statistical Methods: Time series forecasting has been mostly conducted with statistical models such as ARIMA (AutoRegressive Integrated Moving Average (Aryee et al., 2018) and SARIMA (seasonal ARIMA) (Samal, Babu, Das, & Acharaya, 2019). The ARIMA model handles univariate time series and is frequently used in research because of the combined advantage of autoregressive sliding average model. The Multivariate time series is handled with VAR (Vector Autoregression) or VAR-based models such as VARMAX (Jamdade & Jamdade, 2021) and seasonality is incorporated with SARIMAX. (Q. Cheng et al., 2021) proposed a SARIMAX model with external regressors for emergency department occupancy prediction. However, the model showed reduced ability to predict with large time gaps and forecast with better accuracy for one hour predictions as compared to four hours. Recent research incorporates statistical methods with machine learning methods or deep neural networks for improved performance.

3) Machine Learning Methods: In addition to statistical methods which uses a parametric approach, researchers use Machine Learning methods such as Support Vector Regression (Maldonado, Gonzalez, & Crone, 2019), Linear Regression (Ciulla & D'Amico, 2019) and so on. Support Vector Regression with its kernel methods can support non linearity. Linear Regression is usually employed for its simplicity and high interpretability. (Chao, Zhipeng, & Yuanjie, 2019) proposed a SVM based model integrated with cooperative co-evolution algorithm for parametric optimization for the purpose of handling noisy data. Various kernel scale values in SVM was experimented by (Altan & Karasu, 2019) for time series forecasting with volatility estimation. However, these methods are less generalizable across datasets in terms of performance and recent studies have focused on deep learning methods for temporal forecasting. Specialised KNN variants for season-wise forecasting is proposed by (Martínez, Frías, Pérez-Godoy, & Rivera, 2018) which utilises each variant for learning different seasonal trends. This approach reduces erroneous forecasts by targeting cyclical patterns. Support Vector Regression model with polynomial cubic kernel function is proposed by (Beyca, Ervural, Tatoglu, Ozuyar, & Zaim, 2019) for monthly energy consumption forecasting. The model trains multiple input variables including seasonality and population estimates for improved accuracy results. On the other hand, (Kamir, Waldner, & Hochman, 2020) proposed a support vector regression model with radial function for forecasting wheat yields while considering multiple variables such as temperature and rainfall. Support Vector regression is the most commonly used machine learning method for time series analysis and is most often experimented with its kernel function for finding an optimal model for the specific dataset.

4) Graph Neural Networks: Graph Neural Networks (GNN) are used by researchers for modelling multivariate time series owing to its modularity, interpretability and cross modality which makes it suitable for the medical domain (Barbiero, Torné, & Lió, 2021). (Wu et al., 2020) adopted a graph neural network framework with propagation and a dilated inception layer for extracting feature relationships. A multimodal graph neural network combining attention mechanism is proposed by (D. Cheng et al., 2022) for financial time series prediction. However, the model performance is comparable to baselines in capturing market trends despite using the attention module. GNNs are considered to be weak in capturing changes in nodes as this is a sequential process which is deftly managed by recurrent neural networks (RNNs).

5) Recurrent Neural Networks: LSTM (Long short-term memory) networks which are based on RNNs are powerful in learning sequential information and this capability is well

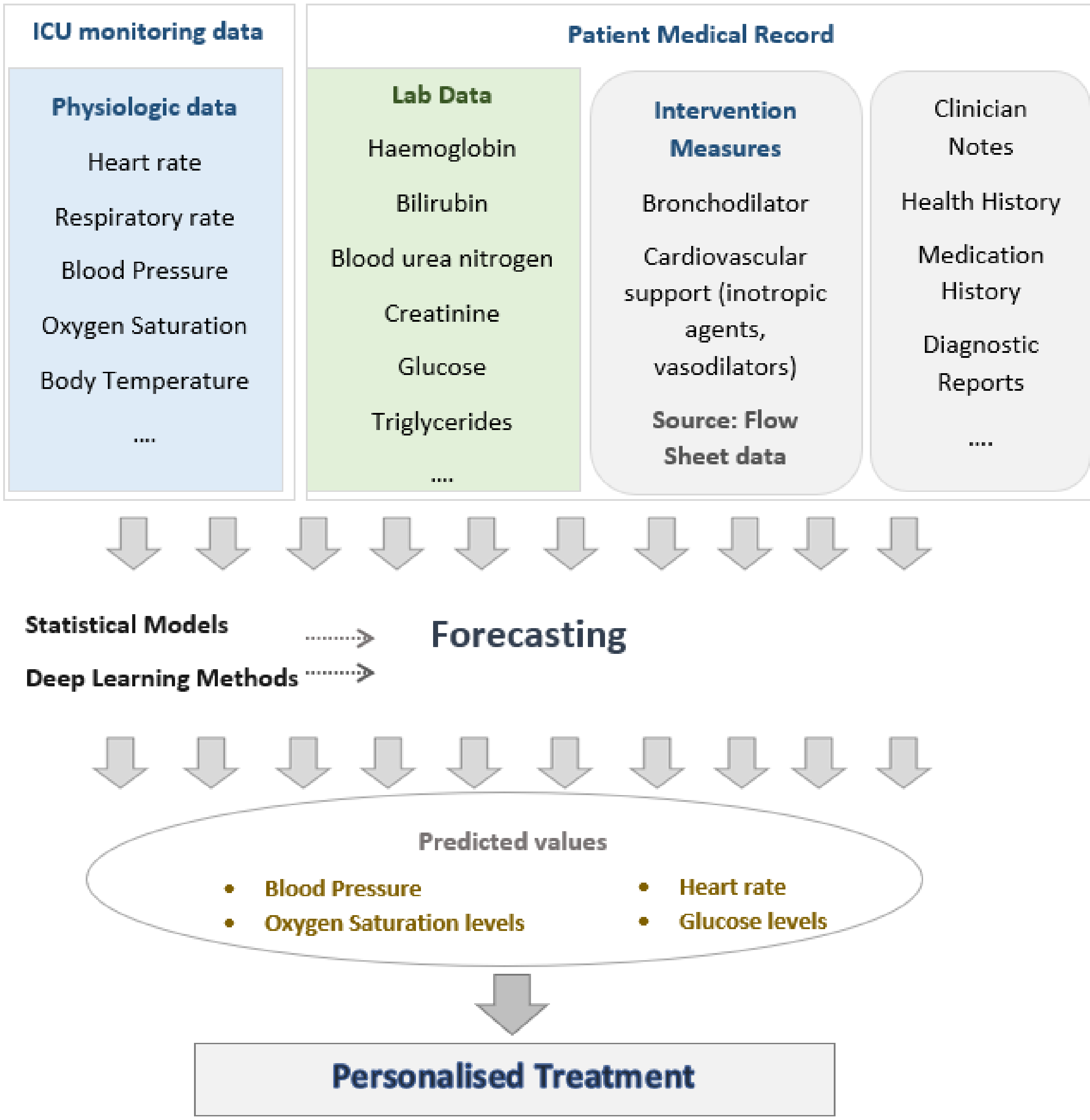

Fig. 2: Personalised medicine in ICU

suited to capture temporal sequences in large volumes of data. Multivariate clinical event forecasting modelled by (J. M. Lee & Hauskrecht, 2021) makes single step predictions with LSTM capturing distant information. Moreover, periodicity is incorporated based on probability distributions. LSTM is useful for making multi-step predictions as demonstrated by the model in (J. Zhang & Nawata, 2018) which formed a six-layered LSTM system for predicting disease outbreaks. (Sagheer & Kotb, 2019) developed a layer wise pre-trained LSTM-based stacked autoencoder as an alternative to weight initialization for multivariate time series forecasting. The approach showed good results when compared to traditional methods, but showed only marginal improvement as compared to baseline LSTM models. A de-noising approach is used by (F. Liu, Cai, Wang, & Lu, 2019) in implementing a LSTM based model with AdaBoost used for weighted distribution in the validation phase for improving prediction results. (S. Huang, Wang, Wu, & Tang, 2019) integrated an autoregressive model with convolutional components and attention mechanism for making revenue predictions on multiple time series. The model performance results were comparable to other LSTM based models, but gave good results in comparison to the traditional models.

GRUs (Gated Recurrent Neural Networks) which is also based on recurrent neural networks are used by researchers for time series modelling. A multi-output forecasting model by (Fox, Ang, Jaiswal, Pop-Busui, & Wiens, 2018) uses GRU coupled with Autoregressive model for forecasting blood glucose levels. The study incorporated sequential dependencies for performance uniformity across data subsets. A two step GRU with imputation and a hidden state decay mechanism is proposed by (Shi et al., 2021) for mortality prediction of ICU patients. The imputation strategies adopted are more generalised and not a customized process for clinical data. LSTMs and variants of LSTM such as stacked LSTM are the most common recurrent neural network based models adopted by time series studies. The long and short term dependencies captured by RNN-based models make it an ideal model for initial experimentation with temporal data.

6) Convolutional Neural Networks: Recurrent Neural Networks are adept at capturing sequence based information, but weak in representing periodicity. Convolutional Neural Networks (CNNs) compensate for this shortcoming and multiple CNNs with residual mapping is employed by (Wan, Mei, Wang, Liu, & Yang, 2019) for making optimal multivariate time series predictions. Convolutional Neural Networks adopt a batch learning process when compared to RNNs' sequential approach. Additionally, CNNs are more suitable for handling missing data in time series. (X. S. Zhang, Tang, Dodge, Zhou, & Wang, 2019) implemented a CNN-based model called MetaPred which accepted a multivariate time series matrix as input into an embedding layer. This layer is passed on to the convolutional layers which output a vector representation which is subsequently passed to three MLP layers for outputting risk probabilities for each patient. The model is observed to perform well for predicting mild cognitive impairment, but not adequate for alzheimers and parkinson's disease prediction. (Ismail, Du, Martinez, & He, 2019) used multichannel CNNs for forecasting severity of Parkinson's disease with a multi-step time series analysis. Prediction results were generated by the model learning a mapping function over each time period. Convolutional layers are efficient in extraction of both the spatial and temporal aspects present in time series data. (Xiao et al., 2021) proposed a convolutional based LSTM model with dual attention layers for effective extraction of the spatio-temporal features. Additionally, the model is also functional at handling the exogenous input features to the temporal data. The Conv-LSTM model proposed by (S. W. Lee & Kim, 2020) efficiently learns from high dimensional time series data and is capable of attaining high predictive power without extensive feature extraction. Additionally, the study used trend sampling for capturing real time data trends. The convolutional layers have the capacity to capture the spatial component which is indirectly associated with temporal data. This capacitates even baseline models to attain high predictive results and is frequently used by researchers in conjunction with RNN-based models such as LSTM.

7) Attention Mechanism: Attention Mechanisms are useful for particularly focussing on certain information which are considered significant in a given problem domain. Attention Mechanisms which are adept at handling sequences have been shown to perform well in time series forecasting. (Eom et al., 2020) used attention mechanism combined with CNN and BiGRU for blood pressure estimation. The performance of attention based model is observed to be higher than the baseline model for single inputs, but similar results are obtained for multiple inputs. (Hu & Zheng, 2020) used a multistage attention network for capturing mutation information from time sequences. As emphasised in the study by (Xiao et al., 2021), adding multiple layers of attention networks contribute to the explicit capturing of interdependent exogenous variables that are relevant in sequential prediction. (Assaf & Schumann, 2019) proposed a CNN model for creating attention-based feature maps with a gradient approach for making multivariate time series predictions. Temporal pattern attention mechanism is employed on CNN by (Shih, Sun, & Lee, 2019) which showed substantial improvement in results as compared to CNN without attention. (F. Liu, Lu, & Cai, 2020) proposed a stacked LSTM model with multi-level attention mechanism, wherein the stacked layers were used as the encoder for extracting the temporal dependencies between the various features. The encoded information from the encoder is transferred to the multi-layered attention module for enhanced predictive performance. A Stacked LSTM model with attention mechanism proposed by (Girkar et al., 2018) is found to attain good prediction results combined with high clinical interpretability in predicting blood pressure in hypotensive patients. However, the study did not consider covariates such as medication, ventilation parameters etc. as input into their model. Attention layers are most often observed to be capable of improving predictive results of temporal data, however, these results are not consistently distributed for all aspects of time series tasks.

8) Generative Adversarial Networks: Generative adversarial networks (GANs) are primarily used in image and video domains. For capturing temporal dependencies in time series data, (Yoon et al., 2019) proposed a time-series version of GAN called TGAN. TGAN uses supervised loss and an embedding network for dimensionality reduction. A hybrid GAN-LSTM is proposed by (Yazdanian & Sharifian, 2021) for forecasting cloud workload assimilating the complexity and volatility of the load traces. The multi-step prediction is performed with a dual layered convolutional network acting as the discriminator. GANs are most commonly used for the purpose of anomaly detection in time series. However, research studies such as (Koochali, Dengel, & Ahmed, 2021) employed a GAN-based model with adverserial training for transforming a deterministic model into a probabilistic model for multivariate time series forecasting.

9) Forecasting Strategies: MIMO (Multiple Input Multiple Output) and DIRMO (Direct Multiple Output) are two forecasting strategies used in time series modelling. The MIMO strategy forecasts values in a single step (Gasparin et al., 2019), however the DIRMO strategy retains dependencies which enables it to perform better in forecast modelling. A comparison of LSTM, BiLSTM and CNN was conducted by (Masum, Chiverton, Liu, & Vuksanovic, 2019) by adopting MIMO (Multiple Input Multiple Output) and DIRMO (Direct Multiple Output) forecasting strategies. The BiLSTM-DIRMO model had better performance than the others in forecasting blood pressure values. However, (Gasparin et al., 2019) employed RNN-based models with Recursive and MIMO forecasting strategies for predicting electricity load, owing to the high computational requirement of the DIRMO strategy. MIMO-based model exhibited good performance results.

The dataset under consideration is significant in deciding the superiority of deep learning methods in time series forecasting.

A comparison of standard machine learning approaches with LSTM for glucose level prediction by (J. Xie & Wang, 2018) indicated that linear regression and support vector regression methods performed better than LSTM. This points to the significance of building LSTM models with approaches such as attention mechanism (X. Zhang et al., 2019) or variants of it (Y. Li, Zhu, Kong, Han, & Zhao, 2019).

10) Hybrid Methods: Statistical methods have been used in conjunction with deep learning to leverage the benefits of both methodologies. (Mathonsi & van Zyl, 2021) presented a multivariate exponential smoothing method with LSTM with results comparable to that of baseline LSTM. However, the coverage of the model was not consistent across groups of observations. (Domingos, de Oliveira, & de Mattos Neto, 2019) proposed a hybrid system with linear models used for series forecasting and a non linear approach for error forecasting. The model uses ARIMA, Multi-layer Perceptron and Support Vector Regression and combines the forecasts for better outcome. A hybrid method with ARIMA and BPNN (back propagation neural network) was proposed by (Hadwan, Al-Maqaleh, Al-Badani, Khan, & Al-Hagery, 2022) for forecasting number of cancer patients. Though the hybrid model resulted in significant error reduction, the BPNN model is more adept at capturing the overall pattern of the series. A ARIMA-ANN hybrid method proposed by (Büyükşahin & Ertekin, 2019) employed time series decomposition followed by merging of the linear and non-linear temporal components for improved time series forecasting. (Caliwag & Lim, 2019) employed hybrid VARMA and LSTM for forecasting multiple cycles with prediction at one cycle ahead that aided in lowering error rates. Building hybrid methods is an optimal approach for combining the linear patterns captured by statistical methods with the non-linear patterns inherent to the time series data for achieving distinguishable predictive results.

11) Ensemble Methods: Many studies use ensemble methods for combining predictions of independent models through a voting mechanism. (D.-R. Liu, Lee, Huang, & Chiu, 2020) proposed a LSTM-Attention layer for predicting air pollution by forecasting PM2.5 concentrations. The predictions provided by the Attention layer is applied with an ensemble method with extreme Gradient Boosting for making secondary predictions. An ensemble method combining decision trees, random forests and gradient boosted trees are proposed by (Galicia, Talavera-Llames, Troncoso, Koprinska, & Martínez-Álvarez, 2019) for multi-step forecasting power consumption. (H. Chen, Guan, & Li, 2021) proposed a multi-factor model integrating LSTM-Attention with XGBoost regression for air quality forecasting. The optimal subtree nodes are used for obtaining the final prediction results. Ensemble methods enable the expression of the most optimal results from various methods. Boosting and bagging methods can lower the prediction errors that standalone methods are prone to when making predictions with temporal data.

Table I summarises some of the research performed in time series forecasting. While forecasting pollutant concentration, (T. Li et al., 2020) demonstrated that multivariate models performed better than univariate methods. Additionally, a hybrid CNN-LSTM has higher predictive power than compared to standalone LSTM model. However, it is observed that the hybrid model is prone to error when a single day's data is input to the model. The error rate is higher when the input length is increased to 14 days. This shows the importance of employing hybrid model rather than a simple architecture such as the model proposed by (Asante, Walker, Seidu, Kpogo, & Zou, 2022), wherein an ARIMA model is observed to be predicting better as compared to a basic LSTM model. It is expected that the performance would be much higher if more layers are added with LSTM as demonstrated by (T. Wang et al., 2020) or a hybrid version such as CNN-LSTM employed by (Parashar et al., 2020). Additionally, with deep learning methods a high performing base layer can be formed, which can be used as a generalised layer and combined with any other layers. (Gu et al., 2020) proposed a generalised model using CNN with dimension reduction, which attains consistent predictive results in conjunction with additional layers such as RNN, GRU or LSTM. This demonstrates the potentiality of deep learning methods when implemented in the most optimal sequence and combination. (Tazarv & Levorato, 2021) predicted blood pressure based on two datasets with CNN-LSTM-MLP layers. The model showed excellent results in predicting blood pressure values from the vital signs dataset. However, the prediction results were similar to other comparable works when predicting with the MIMIC-II dataset. In conjunction with the methodology used, the prediction horizon and sampling period selected plays an important role in improving prediction results as demonstrated by (Song et al., 2021), (T. Wang et al., 2020), (T. Li et al., 2020). A sampling period of 60 seconds is more efficient in forecasting blood pressure levels in a model with LSTM and CNN as demonstrated by (Song et al., 2021). Similarly smaller prediction horizon is observed to give better forecasting estimates for CNN-LSTM model (T. Li et al., 2020) as well as stacked LSTM model (Meng et al., 2020), (T. Wang et al., 2020) indicating the significance of adding more efficient components to the model for predicting long term values.

As outlined in this section, statistical methods, machine learning methods, deep neural networks, hybrid or ensemble methods etc. are systematically used in time series analysis by various studies. While statistical methodology is the most classical approach, more studies are adopting deep neural networks for time series forecasting owing to its high performance. Moreover, recent studies are increasingly using hybrid methods combining statistical and deep neural networks or adding several layers of different variations of neural networks in order to improve predictive power of forecasting models.

IV. Challenges and Issues

Research with ICU data largely relate to predictive modelling of sepsis (Spaeder et al., 2019), cardiac arrests (Matam, Duncan, & Lowe, n.d.), mortality prediction (Ge et al., 2018) and physiologic value forecasting (J. Xie & Wang, 2018). Few of the issues observed in clinical time series modelling are unevenly spaced temporal observations, inconsistent recording of values and missing values. (De Brouwer, Simm, Arany, & Moreau, 2019) adopted a Bayesian update network for

TABLE I: Summary of research in Time Series forecasting with deep learning.

Model	Summary	References
CNN-RNN-AR model	Captures long-term and short-term historical data by handling both linearity and non-linearity	(S. Li et al., 2021)
LSTM-Attention	Weighted Attention mechanisms are adept at handling long term dependencies preserved by sequences captured with LSTM	(Hu & Zheng, 2020), (Yuan et al., 2020), (T. Zhang et al., 2021), (X. Zhang et al., 2019)
Stacked LSTM	Multiple layers of stacked LSTM perform better than standalone LSTM. Performance varies based on the defined prediction horizon	(Meng et al., 2020), (Sagheer & Kotb, 2019), (T. Wang et al., 2020)
Stacked LSTM-ATTN	Temporal attention combined with multiple LSTM capture dependencies effectively and enable multi-step predictions	(Gangopadhyay et al., 2018), (Girkar et al., 2018), (F. Liu et al., 2020)
LSTM-ATTN-XGBoost	Attention mechanisms handle long term dependencies and preliminary predictions are passed to an ensemble learning method	(H. Chen et al., 2021), (D.-R. Liu et al., 2020)
CNN-dimension reduction	Attains interpretable results and allows for consistent interaction between high dimensional features	(Ali et al., 2019), (Gu et al., 2020)
Conv-LSTM	Effective in handling high dimensional time series data. Manages process effectively when output gate is closed	(Essien & Giannetti, 2020), (Indrawan et al., 2021), (S. W. Lee & Kim, 2020), (Shastri et al., 2020)
Conv-LSTM-ATTN	Manages spatial feature extraction effectively for sequenctial data with efficient capturing of long term dependencies	(Singh et al., 2020), (Xiao et al., 2021), (Zheng et al., 2020)
ARIMA-CNN-LSTM	Non-linearity is captured by CNN and LSTM. While Arima captures the linearity, seasonality is incorporated with SARIMA	(Dwivedi et al., 2021), (Ji et al., 2019)
CNN-LSTM	Preserves deterministic feature representations and captures relevant sequencial patterns	(T. Li et al., 2020), (Lin et al., 2017), (Parashar et al., 2020), (H. Xie et al., 2020)
CNN-LSTM-MLP	The combination of forward and backward propagation combined with feature extraction improves the performance of the model	(Phyo & Byun, 2021), (Tazarv & Levorato, 2021)
LSTM-CNN-decomposition	Decomposition of time series data into various frequency spectra aid in compact sequence analysis in tractable sections.	(Asadi & Regan, 2020), (Rezaei et al., 2021), (Song et al., 2021)

processing inconsistent values. This is specifically a challenging factor while handling clinical time series. The data is recorded in uneven time periods, with few of the recordings registered only once per day. Though the lack of data arising due to intermittent recordings cannot be treated as missing data, it remains challenging to address the irregular nature of information. Additionally, the method is less ideal for large volumes of data.

A challenge faced when forecasting with Multivariate Time series is the handling of missing values across multiple features or multiple series. A common approach for handling missing information is imputation, which is more reliable for univariate series. However, for ensuring reliable prediction results it is inevitable to consider global temporal information of the multivariate data. (Tang et al., 2020) attempts a partial solution to the issue by proposing a LSTM based model which generates gradient feedback with discriminator to form a memory module for capturing the local and global temporal dynamics of the data. The global temporal information is fully captured by incorporating adversarial training which contributed to further error reduction of the model.

Time series data owing to its continuous nature is limited in prediction capacity when smaller datasets are involved. Data augmentation is an approach adopted by researchers to combat this limitation, which often involves generating synthetic time series. (Bandara, Hewamalage, Liu, Kang, & Bergmeir, 2021) adopted three techniques for data augmentation viz. GRATIS, moving block bootstrap and dynamic time warping barycentric averaging. The considerable dissimilarity of the newly synthesised series in comparison with the real data is indicative of the contentious performance results.

Additionally, owing to the huge volume of temporal data it is necessary to adopt dimensionality reduction and noise reduction techniques. A dimension reduction strategy employed for forecasting by (Jahangir et al., 2020) is demonstrative of the ability of models to achieve consistent predictive results for all samples. The study also utilised stacked denoising autoencoders to reduce noise typically associated with time series data. The approach is observed to attain high performance results when compared to CNN, LSTM and support vector based methods.

V. Trends and Potential directions

Deep learning approaches such as LSTM, Hybrid methods with GRU and decomposition methods (C. Wang, Liu, Wei, Chen, & Zhang, 2021), MLP, decomposition, ARIMA and SVR (W. Chen, Xu, Chen, & Jiang, 2021), Random Forest with LSTM and decomposition methods (Karijadi & Chou, 2022) are few approaches adopted in recent research whose transferability across domains needs to be studied further. Additionally, introduction of multi-stage attention layers in temporal studies are observed to considerably improve prediction results. (Yin et al., 2021) added internal attention, spatial attention and temporal attention into their proposed model which demonstrated high predictive power than baseline models including an attention based CNN model. Few of the potential directions of the research in time series analysis are addressed here.

Sequence Transformer networks is progressively being utilised for clinical time series research for its ability to capture the various types of invariances specific to the clinical domain. The transformer architecture enable context specific sequence learning and are capable of learning multiple periodic patterns in temporal data. (Zerveas, Jayaraman, Patel, Bhamidipaty, & Eickhoff, 2021) proposed a transformer based model for multivariate time series forecasting with unsupervised learning of the series. A pre-trained transformer model displayed substantial performance capabilities in a unsupervised learning environment. However, the sequence transformer proposed by (Oh, Wang, & Wiens, 2018) is limited to applying transformations uniformly for all features in the series. Consequently, feature-specific learning is restricted leading to exclusion of invariances in the process.

Cross learning from multiple time series is another approach adopted by researchers for efficient information extraction from several variables from various interdependent time series. Feature based and hybrid cross learning methods proposed by (Semenoglou, Spiliotis, Makridakis, & Assimakopoulos, 2021) extracted relevant information from several time series containing varying temporal data. Generalising of cross learning techniques for multiple domains is a deficient process compared to series-wise training. However, in some specialised cases the approach is consistently observed to be attaining high performance results.

In multivariate time series analysis detecting and quantifying correlation among the most significant features is an important step in capturing inter-dependencies between multiple time series or features. While multidimensional recurrence quantification analysis is performed for quantifying the recurrence properties of single time series, multidimensional cross-recurrence quantification analysis proposed by (Wallot, 2019) allows the approach to be utilised for bivariate series. Additionally, the output from cross-recurrence analysis is used for constructing a diagonal cross recurrence profile for extracting time lagged features from multiple time series.

Spiking neurons is an approach adopted in time series research wherein neural networks interact with each other through a series of spikes. The information passed through these spikes, the spike frequency and the distance between the spikes determine the predictive capacity of the models. (Mateńczuk et al., 2021) proposed a standalone spiking neural network (SNN) and a LSTM based spiking neural network (SLSTM) for financial time series forecasting. The optimal number of layers for SNN was estimated to be four, with additional layers not contributing to the performance. The iterative approach adopted by the model resulted in longer training time, but less appreciable performance increase. Based on the study, further research is required to develop an optimised version of spiking neural networks in time series analysis. (Wei, Wang, Niu, & Li, 2021) proposed a convolutional spiking neural network model for extracting temporal features from output generated from the primary gated recurrent neural network layer. The GRU is initially employed on a deconstructed series of temporal data. Grey wolf optimization is combined with the spiking mechanism for determining optimal weights for each of the decomposed time series.

Recent research has adopted various novel approaches for univariate and multivariate time series forecasting. These approaches have manifested good predictive results albeit with limitations. Further studies are required to accentuate the predictive power of the models with additional techniques. While some methods showed dataset-specific or domain-specific performance capabilities, other approaches such as spiking neurons require enhancement with optimisation methodologies to attain replicable results across domains.

VI. Conclusion

The review provides an outline of the methodologies adopted for time series analysis. The conventional methods are based on statistical methods which are linear in nature. However, more research studies are adopting deep learning methods for time series forecasting and classification. The advantage of deep learning methods for time series analysis is that non linear relationships are addressed seamlessly. Additionally, deep learning methods are more adept at making multi-step forecasts, feature engineering, addressing sequential nature of time series data and learning temporal dependencies. Moreover, early studies involved univariate time series analysis. Recent studies have recognised the importance of adopting a multivariate approach in time series analysis for improved performance results. Incorporating of multiple variables or multiple time series in temporal studies have demonstrated high predictive power as compared to single variables.

This work has comprehensively reviewed the various time series analysis methods adopted by researchers including their limitations. This review also provides an indication of challenges faced in the field as well as future trends in time series forecasting, for the development of novel methodologies in temporal analysis. The specific methodology adopted by a researcher is dependent on the research requirement and the domain under study. It is necessary to identify the advantages offered by each method so as to construct a model with a combination of algorithms enabling the model to enhance the predictive potential. Each layer of algorithm is commissioned to handle the different aspects and requirements of the dataset features, thereby qualifying a model with an all encompassing capability for manoeuvring the various facets of the forecasting problem. While many research studies focus on univariate temporal analysis, it is imperative to adopt multivariate analysis for enabling the development of high level real-time temporal forecasting solutions. The review highlights the capability of deep neural networks in handling such complex temporal information for achieving reliable prediction results.

References

Abdollahi, H. (2020). A novel hybrid model for forecasting crude oil price based on time series decomposition. *Applied Energy*, *267*, 115035.

Afrifa-Yamoah, E., Mueller, U. A., Taylor, S., & Fisher, A. (2020). Missing data imputation of high-resolution temporal climate time series data. *Meteorological Applications*, *27*(1), e1873.

Albers, D. J., Levine, M. E., Stuart, A., Mamykina, L., Gluckman, B., & Hripcsak, G. (2018). Mechanistic machine learning: how data assimilation leverages physiologic knowledge using bayesian inference to forecast the future, infer the present, and phenotype. *Journal of the American Medical Informatics Association : JAMIA*, *25*(10), 1392-1401.

Ali, M., Jones, M. W., Xie, X., & Williams, M. (2019). Timecluster: dimension reduction applied to temporal data for visual analytics. *The Visual Computer*, *35*(6), 1013–1026.

Altan, A., & Karasu, S. (2019). The effect of kernel values in support vector machine to forecasting performance of financial time series. *The Journal of Cognitive Systems*, *4*(1), 17–21.

Aryee, G., Kwarteng, E., Essuman, R., Nkansa Agyei, A., Kudzawu, S., Djagbletey, R., … Forson, A. (2018). Estimating the incidence of tuberculosis cases reported at a tertiary hospital in ghana: a time series model approach. *BMC Public Health*, *18*(1), 1–8.

Asadi, R., & Regan, A. C. (2020). A spatio-temporal decomposition based deep neural network for time series forecasting. *Applied Soft Computing*, *87*, 105963.

Asante, D. O., Walker, A. N., Seidu, T. A., Kpogo, S. A., & Zou, J. (2022). Hypertension and diabetes in akatsi south district, ghana: Modeling and forecasting. *BioMed Research International*, *2022*.

Assaf, R., & Schumann, A. (2019). Explainable deep neural networks for multivariate time series predictions. In *Ijcai* (pp. 6488–6490).

Bandara, K., Hewamalage, H., Liu, Y.-H., Kang, Y., & Bergmeir, C. (2021). Improving the accuracy of global forecasting models using time series data augmentation. *Pattern Recognition*, *120*, 108148.

Barbiero, P., Torné, R. V., & Lió, P. (2021). Graph representation forecasting of patient's medical conditions: Toward a digital twin. *Frontiers in genetics*, *12*.

Bedi, J., & Toshniwal, D. (2020). Energy load time-series forecast using decomposition and autoencoder integrated memory network. *Applied Soft Computing*, *93*, 106390.

Beyca, O. F., Ervural, B. C., Tatoglu, E., Ozuyar, P. G., & Zaim, S. (2019). Using machine learning tools for forecasting natural gas consumption in the province of istanbul. *Energy Economics*, *80*, 937–949.

Büyükşahin, Ü. Ç., & Ertekin, Ş. (2019). Improving forecasting accuracy of time series data using a new arima-ann hybrid method and empirical mode decomposition. *Neurocomputing*, *361*, 151–163.

Caliwag, A. C., & Lim, W. (2019). Hybrid varma and lstm method for lithium-ion battery state-of-charge and output voltage forecasting in electric motorcycle applications. *IEEE Access*, *7*, 59680–59689.

Chao, L., Zhipeng, J., & Yuanjie, Z. (2019). A novel reconstructed training-set svm with roulette cooperative coevolution for financial time series classification. *Expert Systems with Applications*, *123*, 283–298.

Chen, H., Guan, M., & Li, H. (2021). Air quality prediction based on integrated dual lstm model. *IEEE Access*, *9*, 93285–93297.

Chen, W., Xu, H., Chen, Z., & Jiang, M. (2021). A novel method for time series prediction based on error decomposition and nonlinear combination of forecasters. *Neurocomputing*, *426*, 85–103.

Cheng, D., Yang, F., Xiang, S., & Liu, J. (2022). Financial time series forecasting with multi-modality graph neural network. *Pattern recognition*, *121*, 108218-.

Cheng, Q., Argon, N. T., Evans, C. S., Liu, Y., Platts-Mills, T. F., & Ziya, S. (2021). Forecasting emergency department hourly occupancy using time series analysis. *The American Journal of Emergency Medicine*, *48*, 177–182.

Ciulla, G., & D'Amico, A. (2019). Building energy performance forecasting: A multiple linear regression approach. *Applied Energy*, *253*, 113500.

De Brouwer, E., Simm, J., Arany, A., & Moreau, Y. (2019). Gru-ode-bayes: Continuous modeling of sporadically-observed time series. *Advances in Neural Information Processing Systems*, *32*.

Domingos, S. d. O., de Oliveira, J. F., & de Mattos Neto, P. S. (2019). An intelligent hybridization of arima with machine learning models for time series forecasting. *Knowledge-Based Systems*, *175*, 72–86.

Dwivedi, S. A., Attry, A., Parekh, D., & Singla, K. (2021). Analysis and forecasting of time-series data using s-arima, cnn and lstm. In *2021 international conference on computing, communication, and intelligent systems (icccis)* (pp. 131–136).

Eom, H., Lee, D., Han, S., Hariyani, Y. S., Lim, Y., Sohn, I., … Park, C. (2020). End-to-end deep learning architecture for continuous blood pressure estimation using attention mechanism. *Sensors*, *20*(8), 2338.

Essien, A., & Giannetti, C. (2020). A deep learning model for smart manufacturing using convolutional lstm neural network autoencoders. *IEEE Transactions on Industrial Informatics*, *16*(9), 6069–6078.

Fox, I., Ang, L., Jaiswal, M., Pop-Busui, R., & Wiens, J. (2018). Deep multi-output forecasting: Learning to accurately predict blood glucose trajectories. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining* (pp. 1387–1395).

Galicia, A., Talavera-Llames, R., Troncoso, A., Koprinska, I., & Martínez-Álvarez, F. (2019). Multi-step forecasting for big data time series based on ensemble learning. *Knowledge-Based Systems*, *163*, 830–841.

Gangopadhyay, T., Tan, S. Y., Huang, G., & Sarkar, S. (2018). Temporal attention and stacked lstms for multivariate time series prediction.

Gasparin, A., Lukovic, S., & Alippi, C. (2019). Deep learning for time series forecasting: The electric load case. *arXiv preprint arXiv:1907.09207*.

Ge, W., Huh, J.-W., Park, Y. R., Lee, J.-H., Kim, Y.-H., & Turchin, A. (2018). An interpretable icu mortality prediction model based on logistic regression and recurrent neural networks with lstm units. *AMIA Annual Symposium proceedings*, *2018*, 460-469.

Girkar, U. M., Uchimido, R., Lehman, L.-W. H., Szolovits, P., Celi, L., & Weng, W.-H. (2018). Predicting blood pressure response to fluid bolus therapy using attention-based neural networks for clinical interpretability. *arXiv preprint arXiv:1812.00699*.

Gu, K., Dang, R., & Prioleau, T. (2020). Neural physiological model: A simple module for blood glucose prediction. In *2020 42nd annual international conference of the ieee engineering in medicine & biology society (embc)* (pp. 5476–5481).

Hadwan, M., Al-Maqaleh, B. M., Al-Badani, F. N., Khan, R. U., & Al-Hagery, M. A. (2022). A hybrid neural network and box-jenkins models for time series forecasting. *CMC-Computers Materials & Continua*, *70*(3), 4829–4845.

Hamidi, O., Borzu, S. R., Maroufizadeh, S., & Amini, P. (2021). Application of multivariate generalized linear mixed model to identify effect of dialysate temperature on physiologic indicators among hemodialysis patients. *Journal of Biostatistics and Epidemiology*, *7*(3), 263–271.

Hu, J., & Zheng, W. (2020). Multistage attention network for multivariate time series prediction. *Neurocomputing*, *383*, 122–137.

Huang, E., Wang, R., Chandrasekaran, U., & Yu, R. (2020). Aortic pressure forecasting with deep learning. In *2020 computing in cardiology* (pp. 1–4).

Huang, S., Wang, D., Wu, X., & Tang, A. (2019). Dsanet: Dual self-attention network for multivariate time series forecasting. In *Proceedings of the 28th acm international conference on information and knowledge management* (pp. 2129–2132).

Indrawan, R., Saadah, S., & Yunanto, P. E. (2021). Blood glucose prediction using convolutional long short-term memory algorithms. *Khazanah Informatika: Jurnal Ilmu Komputer dan Informatika*, *7*(2).

Ismail, N. H., Du, M., Martinez, D., & He, Z. (2019). Multivariate multi-step deep learning time series approach in forecasting parkinson's disease future severity progression. In *Proceedings of the 10th acm international conference on bioinformatics, computational biology and health informatics* (pp. 383–389).

Jahangir, H., Tayarani, H., Baghali, S., Ahmadian, A., Elkamel, A., Golkar, M. A., & Castilla, M. (2020). A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks. *IEEE transactions on industrial informatics*, *16*(4), 2369-2381.

Jamdade, P. G., & Jamdade, S. G. (2021). Modeling and prediction of covid-19 spread in the philippines by october 13, 2020, by using the varmax time series method with preventive measures. *Results in Physics*, *20*, 103694.

Ji, L., Zou, Y., He, K., & Zhu, B. (2019). Carbon futures price forecasting based with arima-cnn-lstm model. *Procedia Computer Science*, *162*, 33–38.

Kamir, E., Waldner, F., & Hochman, Z. (2020). Estimating wheat yields in australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, *160*, 124–135.

Karijadi, I., & Chou, S.-Y. (2022). A hybrid rf-lstm based on ceemdan for improving the accuracy of building energy consumption prediction. *Energy and Buildings*, *259*, 111908.

Karim, F., Majumdar, S., Darabi, H., & Harford, S. (2019). Multivariate lstm-fcns for time series classification. *Neural Networks*, *116*, 237–245.

Koochali, A., Dengel, A., & Ahmed, S. (2021). If you like it, gan it—probabilistic multivariate times series forecast with gan. In *Engineering proceedings* (Vol. 5, p. 40).

Lee, J. M., & Hauskrecht, M. (2021). Modeling multivariate clinical event time-series with recurrent temporal mechanisms. *Artificial Intelligence in Medicine*, *112*, 102021.

Lee, S. W., & Kim, H. Y. (2020). Stock market forecasting with super-high dimensional time-series data using convlstm, trend sampling, and specialized data augmentation. *Expert Systems with Applications*, *161*, 113704.

Li, S., Huang, H., & Lu, W. (2021). A neural networks based method for multivariate time-series forecasting. *IEEE Access*, *9*, 63915–63924.

Li, T., Hua, M., & Wu, X. (2020). A hybrid cnn-lstm model for forecasting particulate matter (pm2. 5). *IEEE Access*, *8*, 26933–26940.

Li, Y., Zhu, Z., Kong, D., Han, H., & Zhao, Y. (2019). Ea-lstm: Evolutionary attention-based lstm for time series prediction. *Knowledge-Based Systems*, *181*, 104785.

Lin, T., Guo, T., & Aberer, K. (2017). Hybrid neural networks for learning the trend in time series. In *Proceedings of the twenty-sixth international joint conference on artificial intelligence* (pp. 2273–2279).

Liu, D.-R., Lee, S.-J., Huang, Y., & Chiu, C.-J. (2020). Air pollution forecasting based on attention-based lstm neural network and ensemble learning. *Expert Systems*, *37*(3), e12511.

Liu, F., Cai, M., Wang, L., & Lu, Y. (2019). An ensemble model based on adaptive noise reducer and over-fitting prevention lstm for multivariate time series forecasting. *IEEE Access*, *7*, 26102–26115.

Liu, F., Lu, Y., & Cai, M. (2020). A hybrid method with adaptive sub-series clustering and attention-based stacked residual lstms for multivariate time series forecasting. *IEEE Access*, *8*, 62423–62438.

Maldonado, S., Gonzalez, A., & Crone, S. (2019). Automatic time series analysis for electric load forecasting via support vector regression. *Applied Soft Computing*, *83*, 105616.

Martínez, F., Frías, M. P., Pérez-Godoy, M. D., & Rivera, A. J. (2018). Dealing with seasonality by narrowing the training set in time series forecasting with knn. *Expert systems with applications*, *103*, 38–48.

Masum, S., Chiverton, J. P., Liu, Y., & Vuksanovic, B. (2019). Investigation of machine learning techniques in forecasting of blood pressure time series data. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 269–282).

Matam, B. R., Duncan, H., & Lowe, D. (n.d.). Machine learning based framework to predict cardiac arrests in a paediatric intensive care unit: Prediction of cardiac arrests. *Journal of clinical monitoring and computing*, *33*(4), 713-724.

Mateńczuk, K., Kozina, A., Markowska, A., Czerniachowska, K., Kaczmarczyk, K., Golec, P., … others (2021). Financial time series forecasting: Comparison of traditional and spiking neural networks. *Procedia Computer Science*, *192*, 5023–5029.

Mathonsi, T., & van Zyl, T. L. (2021). A statistics and deep learning hybrid method for multivariate time series forecasting and mortality modeling. *Forecasting*, *4*(1), 1–25.

Meng, X., Liu, M., & Wu, Q. (2020). Prediction of rice yield via stacked lstm. *International Journal of Agricultural and Environmental Information Systems (IJAEIS)*, *11*(1), 86–95.

Oh, J., Wang, J., & Wiens, J. (2018). Learning to exploit invariances in clinical time-series data using sequence transformer networks. In *Machine learning for healthcare conference* (pp. 332–347).

Oyeleye, M., Chen, T., Titarenko, S., & Antoniou, G. (2022). A predictive analysis of heart rates using machine learning techniques. *International Journal of Environmental Research and Public Health*, *19*(4), 2417.

Parashar, A., Mohan, Y., & Rathee, N. (2020). *Forecasting covid-19 cases in india using deep cnn lstm model* (Tech. Rep.). EasyChair.

Phyo, P. P., & Byun, Y.-C. (2021). Hybrid ensemble deep learning-based approach for time series energy prediction. *Symmetry*, *13*(10), 1942.

Rezaei, H., Faaljou, H., & Mansourfar, G. (2021). Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, *169*, 114332.

Sagheer, A., & Kotb, M. (2019). Unsupervised pre-training of a deep lstm-based stacked autoencoder for multivariate time series forecasting problems. *Scientific reports*, *9*(1), 1–16.

Samal, K. K. R., Babu, K. S., Das, S. K., & Acharaya, A. (2019). Time series based air pollution forecasting using sarima and prophet model. In *proceedings of the 2019 international conference on information technology and computer communications* (pp. 80–85).

Seid, S. S., Ibro, S. A., Ahmed, A. A., Akuma, A. O., Reta, E. Y., Haso, T. K., & Fata, G. A. (2019). Causes and factors associated with neonatal mortality in neonatal intensive care unit (nicu) of jimma university medical center, jimma, south west ethiopia. *Pediatric health, medicine and therapeutics*, *10*, 39.

Semenoglou, A.-A., Spiliotis, E., Makridakis, S., & Assimakopoulos, V. (2021). Investigating the accuracy of cross-learning time series forecasting methods. *International Journal of Forecasting*, *37*(3), 1072–1084.

Shastri, S., Singh, K., Kumar, S., Kour, P., & Mansotra, V. (2020). Time series forecasting of covid-19 using deep learning models: India-usa comparative case study. *Chaos, Solitons & Fractals*, *140*, 110227.

Shi, Z., Wang, S., Yue, L., Pang, L., Zuo, X., Zuo, W., & Li, X. (2021). Deep dynamic imputation of clinical time series for mortality prediction. *Information Sciences*, *579*, 607–622.

Shih, S.-Y., Sun, F.-K., & Lee, H.-y. (2019). Temporal pattern attention for multivariate time series forecasting. *Machine Learning*, *108*(8), 1421–1441.

Singh, S. P., Sharma, M. K., Lay-Ekuakille, A., Gangwar, D., & Gupta, S. (2020). Deep convlstm with self-attention for human activity decoding using wearable sensors. *IEEE Sensors Journal*, *21*(6), 8575–8582.

Song, X., Zhu, L., Feng, X., Wu, H., & Li, Y. (2021). Combined forecast model of lstm-cnn hypertension based on eemd. In *2021 4th international conference on signal processing and machine learning* (pp. 117–122).

Spaeder, M. C., Moorman, J. R., Tran, C. A., Keim-Malpass, J., Zschaebitz, J. V., Lake, D. E., & Clark, M. T. (2019). Predictive analytics in the pediatric intensive care unit for early identification of sepsis: capturing the context of age. *Pediatric research*, *86*(5), 655-661.

Staffini, A., Svensson, T., Chung, U.-i., & Svensson, A. K. (2021). Heart rate modeling and prediction using autoregressive models and deep learning. *Sensors*, *22*(1), 34.

Tang, X., Yao, H., Sun, Y., Aggarwal, C., Mitra, P., & Wang, S. (2020). Joint modeling of local and global temporal dynamics for multivariate time series forecasting with missing values. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 34, pp. 5956–5963).

Tazarv, A., & Levorato, M. (2021). A deep learning approach to predict blood pressure from ppg signals. In *2021 43rd annual international conference of the ieee engineering in medicine & biology society (embc)* (pp. 5658–5662).

Thao, P. T. N., Tra, T. T., Son, N. T., & Wada, K. (2018). Reduction in the il-6 level at 24 h after admission to the intensive care unit is a survival predictor for vietnamese patients with sepsis and septic shock: a prospective study. *BMC Emergency Medicine*, *18*(1), 1–7.

Wallot, S. (2019). Multidimensional cross-recurrence quantification analysis (mdcrqa)–a method for quantifying correlation between multivariate time-series. *Multivariate behavioral research*, *54*(2), 173–191.

Wan, R., Mei, S., Wang, J., Liu, M., & Yang, F. (2019). Multivariate temporal convolutional network: A deep neural networks approach for multivariate time series forecasting. *Electronics*, *8*(8).

Wang, C., Liu, Z., Wei, H., Chen, L., & Zhang, H. (2021). Hybrid deep learning model for short-term wind speed forecasting based on time series decomposition and gated recurrent unit. *Complex System Modeling and Simulation*, *1*(4), 308–321.

Wang, T., Li, W., & Lewis, D. (2020). Blood glucose forecasting using lstm variants under the context of open source artificial pancreas system.

Wei, D., Wang, J., Niu, X., & Li, Z. (2021). Wind speed forecasting system based on gated recurrent units and convolutional spiking neural networks. *Applied Energy*, *292*, 116842.

Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020). Connecting the dots: Multivariate time series forecasting with graph neural networks. In *Proceedings of the 26th acm sigkdd international conference on knowledge discovery & data mining* (pp. 753–763).

Xiao, Y., Yin, H., Zhang, Y., Qi, H., Zhang, Y., & Liu, Z. (2021). A dual-stage attention-based conv-lstm network for spatio-temporal correlation and multivariate time series prediction. *International Journal of Intelligent Systems*, *36*(5), 2036–2057.

Xie, H., Zhang, L., & Lim, C. P. (2020). Evolving cnn-lstm models for time series prediction using enhanced grey wolf optimizer. *IEEE Access*, *8*, 161519–161541.

Xie, J., & Wang, Q. (2018). Benchmark machine learning approaches with classical time series approaches on the blood glucose level prediction challenge. In *Khd@ ijcai.*

Yazdanian, P., & Sharifian, S. (2021). E2lg: a multiscale ensemble of lstm/gan deep learning architecture for multistep-ahead cloud workload prediction. *The Journal of Supercomputing*, *77*(10), 11052–11082.

Yee, C. R., Narain, N. R., Akmaev, V. R., & Vemulapalli, V. (2019). A data-driven approach to predicting septic shock in the intensive care unit. *Biomedical informatics insights*, *11*, 1178222619885147.

Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2021). Multi-stage attention spatial-temporal graph networks for traffic prediction. *Neurocomputing*, *428*, 42–53.

Yoon, J. H., Mu, L., Chen, L., Dubrawski, A., Hravnak, M., Pinsky, M. R., & Clermont, G. (2019). Predicting tachycardia as a surrogate for instability in the intensive care unit. *Journal of Clinical Monitoring and Computing*, *33*(6), 973–985.

Yuan, Y., Lin, L., Huo, L.-Z., Kong, Y.-L., Zhou, Z.-G., Wu, B., & Jia, Y. (2020). Using an attention-based lstm encoder–decoder network for near real-time disturbance detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 1819–1832.

Zerveas, G., Jayaraman, S., Patel, D., Bhamidipaty, A., & Eickhoff, C. (2021). A transformer-based framework for multivariate time series representation learning. In *Proceedings of the 27th acm sigkdd conference on knowledge discovery & data mining* (pp. 2114–2124).

Zhang, J., & Nawata, K. (2018). Multi-step prediction for influenza outbreak by an adjusted long short-term memory. *Epidemiology & Infection*, *146*(7), 809–816.

Zhang, T., Zheng, X.-Q., & Liu, M.-X. (2021). Multiscale attention-based lstm for ship motion prediction. *Ocean Engineering*, *230*, 109066.

Zhang, X., Liang, X., Zhiyuli, A., Zhang, S., Xu, R., & Wu, B. (2019). At-lstm: An attention-based lstm model for financial time series prediction. In *Iop conference series: Materials science and engineering* (Vol. 569, p. 052037).

Zhang, X. S., Tang, F., Dodge, H. H., Zhou, J., & Wang, F. (2019). Metapred: Meta-learning for clinical risk prediction with limited patient electronic health records. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining* (pp. 2487–2495).

Zheng, H., Lin, F., Feng, X., & Chen, Y. (2020). A hybrid deep learning model with attention-based conv-lstm networks for short-term traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *22*(11), 6910–6920.