Worldwide per-capita caffeine consumption time series prediction using CNN and baseline evaluation.

Halah Ahmad Abd Almeneem Mhamd [1,2,*]

1 is with the Department of Mathematics, Darb University College, Jazan University, Jazan, 45142, Saudi Arabia.
2 On leave from the Department of Mathematics, Faculty of Science, Minia University, Minia, Egypt.

0000-0001-8690-0592

DOI: https://doi.org/10.56293/IJASR.2024.5904

IJASR 2024
VOLUME 7`
ISSUE 3 MAY – JUNE **ISSN: 2581-7876**

Abstract: It is undeniable that high temperatures, especially those that are above 35 degrees Celsius, make people feel "uncomfortable," which impairs their ability to work and think clearly, makes them irritable, hinders their capacity to attain their typical productivity, and makes them frustrated. One turns to what drives him to work—coffee, tea, and chocolate —to get over this dissatisfaction. Therefore, predicting per capita of coffee and caffeine consumption is essential and requires the identification of powerful and effective methodologies capable of deducing random dependency between the past and the future and creating an accurate prediction of future behavior. In this study, we use deep learning models for time series prediction with multiple steps ahead and evaluate their effectiveness, in forecasting the three-time series; KSA per-capita coffee consumption data from 1961 to 2019, KSA Percentage distribution of Employment-to-Population Ratio from Age of 15 Years Old from the Second Quarter (Q2) in the year 2016 to Second Quarter in the year 2022, and Average temperatures in some province from north to south of Saudi Arabia. Realizing the direction of the curve of these data statistics through extracting effective features in prediction by learning from the raw inputs. Experimental analysis shows that the suggested one-dimensional approaches; the convolution neural network 1D-CNN, convolution Long Short-Term Memory networks or 1D-ConvLSTM, and Encoder-Decoder LSTM or ED-LSTM can predict future values of coffee per capita consumption with high predictive accuracy.

Keywords: Per-capita Coffee Consumers, Multi-Step prediction, 1D-CNN, MSE.

1. Introduction

Climate change has recently resulted in a rise in the frequency of natural hazards, which only threaten life and property when they occur in populated or inhabited areas. Depending on the geographic zone, one can find: 1) Floods typically occur when there is heavy rain linked with strong thunderstorms, hurricanes, tropical storms intensified by melted ice (a result of global warming), and glacier water or snow water pouring across ice sheets or snowfields, according to [1][2] it manifests in low-lying geomorphological regions. Extreme rainfall is the primary cause of more severe floods in countries, along with the soil's decreased water-holding ability. In order to combat flood events, Mohamed Fofana et al. [1] identified two categories of methods: structural methods, which are ineffective at forecasting impending floods and cannot warn the populace with any lead time in advance, and non-structural methods, which permit prediction with a reasonable lead time. Since the world is advancing towards "smart" technology and in the ICT Era, Nkordeh et al [3] investigated that the Internet of Things (IoT) can be employed in the management and control of environmental hazards. They studied how satellites could be used to regulate several types of environmental hazards, such as pollution, ineffective waste disposal, and natural disasters, 2) the presence and spread of some illnesses may differ. Akinbobola and Hamisu [4] explored that climatic and environmental factors might affect the prevalence of malaria in different parts of Nigeria, 3) The consequence of both moisture and temperature on the production and cultivation of some agricultural crops, such as coffee, and 4) Weather and climate conditions have a direct impact on human health and well-being; the milder the weather, the more stable human life will be in terms of health and practicality. Enjoyment of cultural rights, including cultural activities, significant venues for cultural interactions, and ways of life, are threatened by climate change. Recently, United Nations Human Rights stated that successful solutions to climate change will necessitate modifications to global lifestyles, including those related to production, consumption, and mobility, to mention a few. Culture, science, innovation, and the assertion of cultural rights will play a key role in these changes. Through a history rich in rituals and traditions, the concepts of generosity and hospitality, and the human, aesthetic, and artistic presence in songs, poems, and paintings, coffee has come to be identified with the cultural heritage of the Kingdom of Saudi Arabia. It became a crucial component of Saudi culture and folk heritage, and a cultural mark that distinguishes the Kingdom, whether through its cultivation, or methods of preparation; and presentation to guests out of the high status of this cultural and national symbol, the year 2022 AD was named the "Year of Saudi Coffee". All segments of society prefer to drink coffee anytime and anywhere. Coffee is produced in more than 50 countries around the world, South America and Asia dominate global production, while Europe accounts for the bulk of global consumption. While Brazil leads the world's coffee-producing and exporting countries, Finland tops the list of the world's most consuming people per capita, and the United States in terms of total consumption, for more details one can find in [5][6]. Figure 1 shows the countries that consume the most coffee in the world. As recorded in [7] Saudi Arabia's per-capita consumption of coffee reached 2.16 kg in 2020. The Saudi population has increased by 1.2% [8] [9]. Because of the importance of coffee to rural economies in many tropical countries, the International Coffee Council (IPCC) [10][11] examined the potential impacts of a warmer climate on coffee production in the Americas and Africa. According to [12] the weather temperature in KSA is approximately normally distributed. Dark drinks such as coffee, tea, and chocolate are mostly caffeine [13]. Arabic coffee, or "gahwa," is a popular hot beverage offered every day at all local social occasions and gatherings in Saudi Arabia. "Gahwa" is typically eaten with dates and sweets and is served in tiny cups. Salwa Ali Albar, et al [14] found that more research is needed to expand existing evidence and determine the effects of timing and dose of daily habitual caffeine consumption among adults living freely with diabetes. Several traditional drinks are being replaced by coffee as a staple of the contemporary lifestyle in Saudi Arabia. Several people in the nation are searching for novel coffee flavors and fresh flavors to supplement conventional designs [15]. Major cities, such as London, Paris, New York, and Beijing, are famous for their cafes, so you find their own flavor and taste in their distinguished cafes, not only hot coffee, but extended to iced and fruity coffee, even coffee no longer only black, but has taken multiple colors.

The goal of prediction is to predict future behavior. Without knowledge of basic mechanics, forecasting makes it possible to determine the optimal action to be taken.

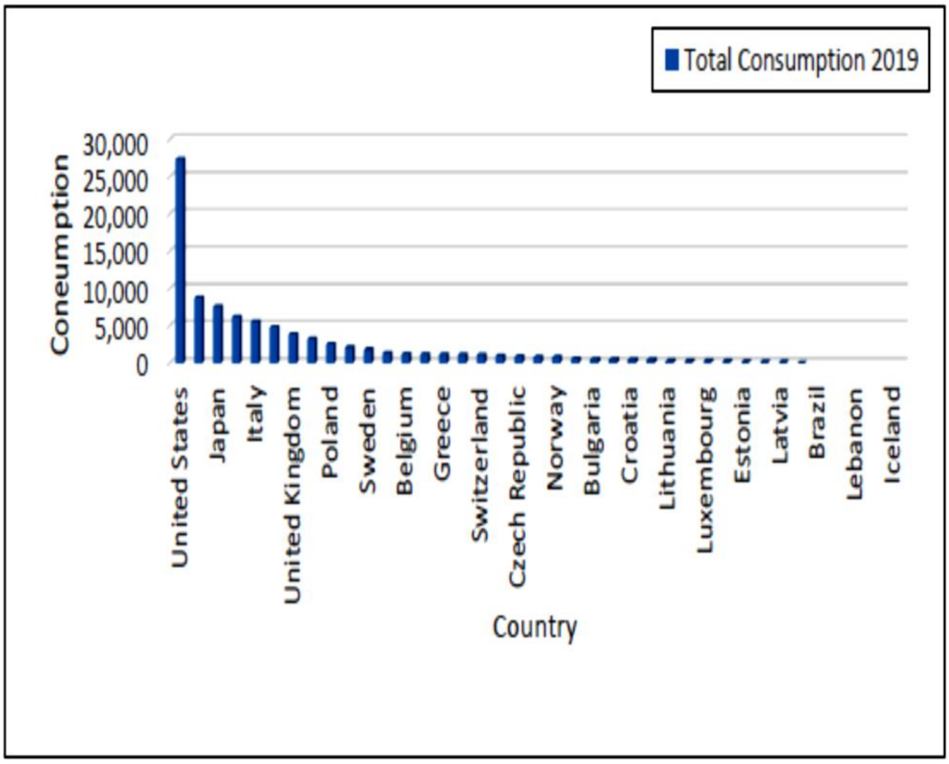

Figure 1. The countries that consume the most coffee in the world in the year 2019.

The observed data measured with real values at regular intervals model is referred to as time series data. Time series data is the best way to predict, so it is widely used in many scientific fields, and commercial applications and has a number of significant real-world applications, including biometric authentication such as online signature verification, analysis in medicine, stock prediction in finance, energy usage prediction in power grids, human activity recognition [16] [17], weather prediction in meteorology, and solar activity prediction in space weather.

Through the application of a variety of time series analysis tools, including statistical techniques, machine learning (ML), and deep learning, researchers have recently developed the ability to reduce time series features into manageable scales, which can then be used to predict how a system will change over time [18] [19]. A well-known deep learning model is the Convolutional neural network (CNN). The primary characteristic of CNNs is their ability to conduct several processing units (such as convolution, pooling, sigmoid/hyperbolic tangent squashing, rectifier, and normalizing) alternately. Authors [20]-[23] developed a modified version of 2D CNNs called 1D Convolutional Neural, the better choice for processing 1D Data values in a variety of applications. Today experts and academics have now conducted extensive research on the multi-step prediction of time series [24]. A multi-step time series prediction can forecast several time steps while a one-step time series prediction can only anticipate the value of the following time step. Cortez et al [25] presented multi-step time series prediction intervals using Neural Computing. To learn the dynamic changes of various scales, Gao et al [26] suggested a framework to implement multi-step prediction using a number of self-attention blocks. The paper is structured as follows: A background and overview of related publications are presented in the introduction section. Section 2 presents some statistical analysis. Section 3 presents the details of the materials and the applied methods, and Section 4 presents experiments. Section 5 provides an evaluation and discussion. Section 6 concludes the paper with future work.

Problem definition

Climate change threatens the enjoyment of cultural rights and ways of life. Given recent data sequences, what is the expected data for the years ahead? To accomplish this, a predictive model must project the total data for each year over the following years. That is; using the prior independent variables as a guide, forecast the dependent variables. Because there are multiple forecast steps, this issue formulation is technically known as a multi-step time series forecasting problem. A univariate multi-step time series forecasting model is one that only uses one input variable as its input.

Time series forecasting

Time series forecasting actually entails calculating the time series value in the interval marked (t) based on the previously known interval denoted (t-1). There are two categories of time series forecasting models: (i) parametric or probabilistic models (such as exponential smoothing and the popular random models Autoregressive (AR), Moving Averaging (MA), and ARIMA), and (ii) non-parametric or computational models (such as artificial neural networks (ANN), support vector machines (SVM) and nearest neighbors).

The exponential smoothing approach model has been improved for several years [27]. Numerous well-known time series prediction algorithms are built on this foundation. The neural networks (NN) can learn a variety of linear and nonlinear data without the requirement for in-depth knowledge of the data or assumptions, time series data are transformed into a supervised learning structure using a sliding window. Lecun, Y. et al [28] introduced CNNs, a well-known deep-learning architecture that originally came from mammals' native visual systems. Wang, H.-z. et al [29] investigated that CNN generated highly promising results; when used for time series prediction.

Tran et al [30] summarized multi-step forecasting methods into five possible forecasting strategies: Recursive, Direct, DirRec, MIMO, and DIRMO strategies. Figure 2 shows the various forecasting techniques with connections showing how they are related. The multiple output MO technique entails creating a single model that is able to accurately anticipate the whole forecast sequence in a single prediction.

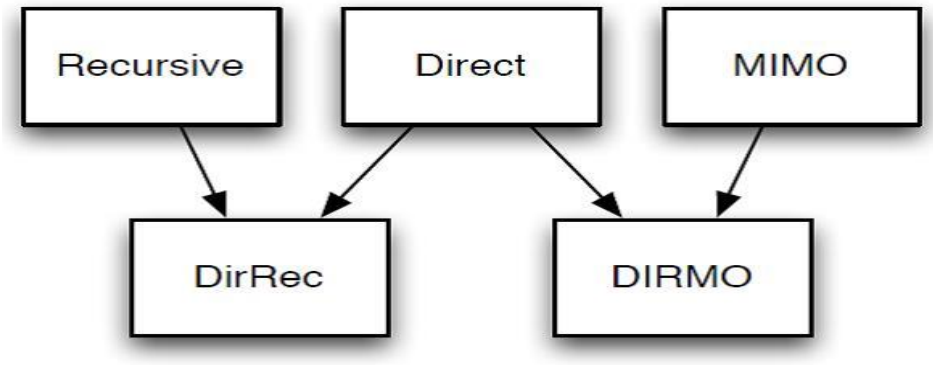

Figure 2. The various forecasting techniques with connections show how they are related. Copied from [30]

Convolution Neural Network (CNN)

CNNs learn spatial hierarchies of features by using multiple building blocks, such as convolution, pooling layers, and fully connected layers [31][32]. The terms "deep learning" and "neural networks" are frequently used interchangeably. It's important to know that the "deep" in deep learning just denotes the number of layers in a neural network. CNN is a subset of deep learning. CNNs are now receiving more interest for time series forecasting. The combination of recursive and direct strategies [33] has become popular like many support vector regression SVR models that were separately trained using the same training data [34]. Grigorievskiy et al [35] used the recursive and the direct strategy in Long-term time series prediction and concluded that the combination of the two strategies gives better results. Ben Taieb, S., Sorjamaa, A., and Bontempi, G. [36] introduced a review of single-output versus multiple-output methods showing the direct technique more effective option than the recursive technique. Persistent approaches have been applied in recent years to the analysis of data and dynamic systems. By employing these methods, it is possible to gain crucial knowledge about the periodicity, bi-stability, and chaos of the underlying systems by analyzing the shape and patterns of the data and systems. Many researchers characterized the simplest persistence model as:

$$\widehat{y_{i+k_i}} = y_i \qquad (1)$$

They assumed that the output would remain constant in the future which is known as simple or naive persistence; also used it as a benchmark for the outcomes of novel models and added a subsection persistence model to a set of evaluation criteria including the Mean Absolute Error and Root Mean Square Error. In our experimental evaluation, we used this as an assessed reference.

Long short-term memory (LSTM)

Long short-term memory is a deep learning technique known as LSTM that substitutes conventional neural networks with three gates. It is a specific type of recurrent neural network (RNN). LSTM is incredibly effective at solving a wide range of issues and is currently popular for time-series issues. By storing information for a long period, LSTM is specifically designed to solve difficulties with long-term dependence, more about LSTM one can find in [37]. Both the input and the output in a multi-step series prediction have varying lengths. Encoding Decoding LSTM, also known as ED-LSTM networks, is used to manage input and output sequences of varying length by first encoding each sequence individually using a latent vector representation, and then decoding the individual sequences from the representation. Given an input sequence, the ED-LSTM computes a sequence of hidden states during the encoding stage. It establishes a distribution over the output sequence given the input sequence during the decoding phase.

2. Statistical analysis

Prediction in statistics is a subset of statistical inference. Statistics, which is not always the same as prediction over time, can be described as a way of transmitting knowledge about a sample of a population to the entire population and to other related populations. In order to describe the data pattern of KSA per-capita coffee consumption, KSA total average temperature, and KSA Population Working Ratio from the year 1996 until the year 2020, we first generated various statistical measures such as the mean, kurtosis, standard deviation (STD), and skew. Table 1 shows that the kurtosis is less than 3, and the distribution will have a smaller tail or fewer outliers than expected. Compared to normal distribution it will have a wider bell-shaped distribution and a lower peak. One of the most used metrics in data science is covariance. Understanding covariance in all of its complexities opens a lot of possibilities for comprehending multivariate data. From the data displayed in Table 2, covariance is discovered to be positive. It indicates that variables vary in the same way, i.e.; there is a positive relationship between them. We also calculated the Pearson correlation coefficient [38] in order to evaluate the linear relationship between each pair of data sets, Table 3 shows the correlation data, all of which are positive. As a result, we believe that when applying our predictive models to the per-capita coffee consumer, the average KSA Population Working Ratio and the total average temperature may be significant influences that shouldn't be overlooked.

Table 1 Skew, Kurtosis, Mean, and Standard deviation of the used data from the year 1996 until the year 2020.

Data Set Measures	A KSA per-capita coffee consumer	B Average KSA Population Working Ratio	C KSA Total average temperature
Skew	-0.393	0.424	-0.101
Kurtosis	-1.059	-1.380	-1.367
Mean	1.525	39.433	25.506
Standard deviation	0.280	3.511	5.170

Table 2 covariance of each variable.

	KSA Total average temperature	KSA per-capita coffee consumer	Average KSA Population Working Ratio
KSA Total average temperature	29.15340833	0.48222408	0.60424242
KSA per-capita coffee consumer	0.48222408	0.08536713	0.82687758
Average KSA Population Working Ratio	0.60424242	0.82687758	13.44424242

Table 3 person correlations.

	KSA per-capita coffee consumer	Average KSA Population Working Ratio	KSA Total average temperature
KSA per-capita coffee consumer	1	0.772	0.306
Average KSA Population Working Ratio	0.772	1	0.067
KSA Total average temperature	0.306	0.067	1

3. Materials and Methods

Methods

Applications and the study of phenomena inspired the proposal of new random processes. Using deep learning models for time series prediction has received a lot of interest in light of the recent deep learning revolution. In this paper, we propose to apply the deep learning approaches based on Multi-step forecasting; the one-dimensional Convolutional neural networks (1D-CNN), one of the non-parametric or computational models, and the encoder-decoder LSTM multi-step forecasting model. We use (individually) the statistical metrics mean square error (MSE) and root mean square error (RMSE), to evaluate the outcomes. They are described in (2) and (3).

$$MSE = \frac{1}{N}\sum_{t=1}^{N}[\hat{y}(t) - y(t)]^2 \qquad (2)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}[\hat{y}(t) - y(t)]^2} \qquad (3)$$

Data Resources

The data we analyzed statistically in paragraph 3 were collected from individual sources. We predict three data statistics time series. We predict three data statistics time series. The first time series data are modeled from statistics around the world of Helgi library statistics to directly monitor coffee consumption per capita [7]. As the age group over the age of 15 is most likely to consume coffee the prediction is based on Saudi Arabian General Authority for Statistics web data [8][9] of the percentage distribution of the employment to population ratio for those people where the second time series data are modeled from. Figure 3 shows the Percentage distribution of this data. According to [12] Figure 4 shows the average temperature of a number of cities of different regions in KSA over 12 months within years from 1991 to 2020. Raw weather data one can obtain from the Visual Crossing website [39].

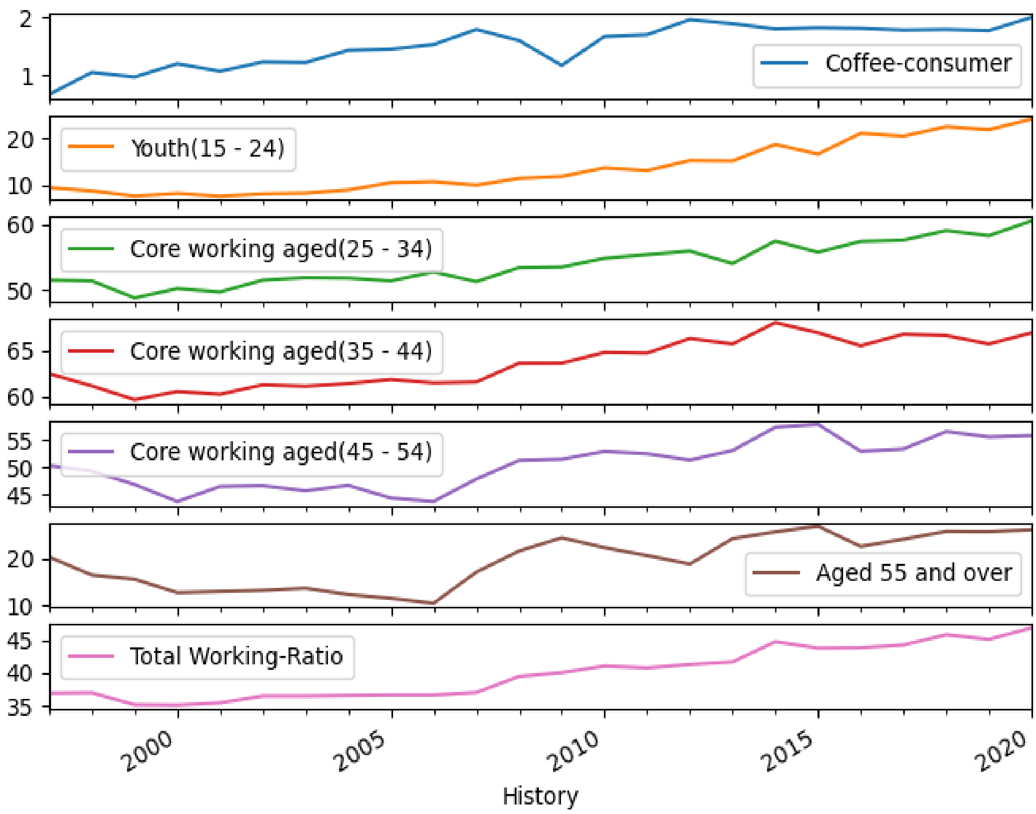

Figure 3. Per-capeta coffee consumer and Percentage distribution of Employment to Population Ratio from 15 aged to over 55 Years Old in Saudi Arabia. The period from Second Quarter (Q2) in the year 2016 to the Second Quarter in the year 2022.

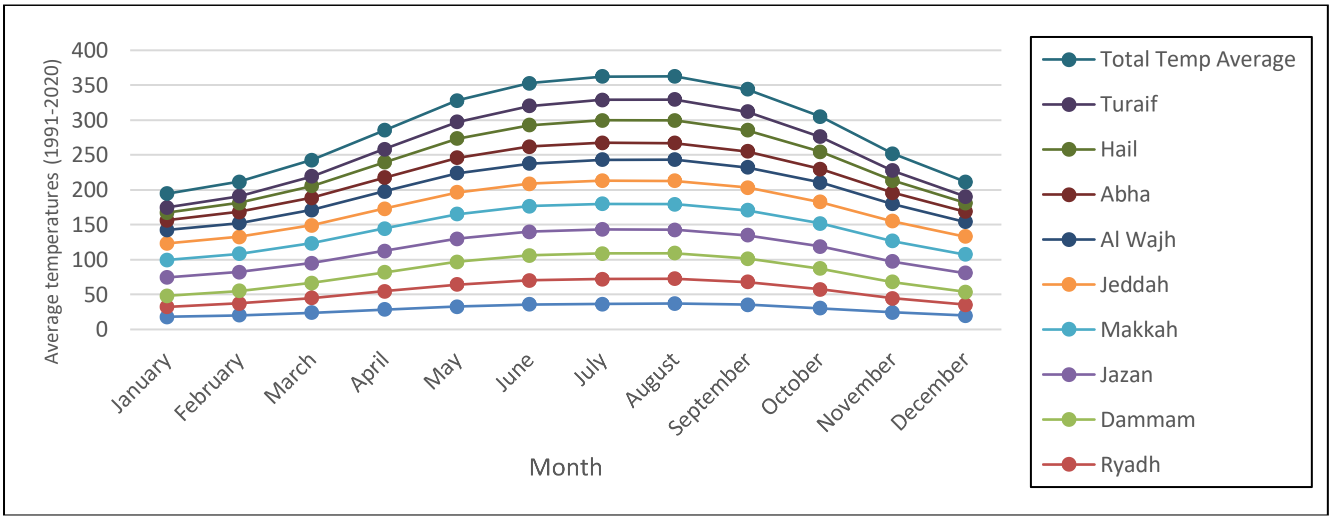

Figure 4. Average temperature of a number of cities in KSA over 12 months within years from 1991 to 2020 and the top layer represents the total temperature average.

4. Experiment

Machine learning or ML algorithms deal with learning challenging and sophisticated problems like time-series forecasting since they are based on statistical analysis approaches to recognize patterns from a set of data. Due to its design and inspiration from how human brains function, deep learning or DL algorithms are ideally suited to address this issue because they may use several customized hyperparameters to extract hidden patterns from features. Figure 5 shows the flowchart of our proposed approaches.

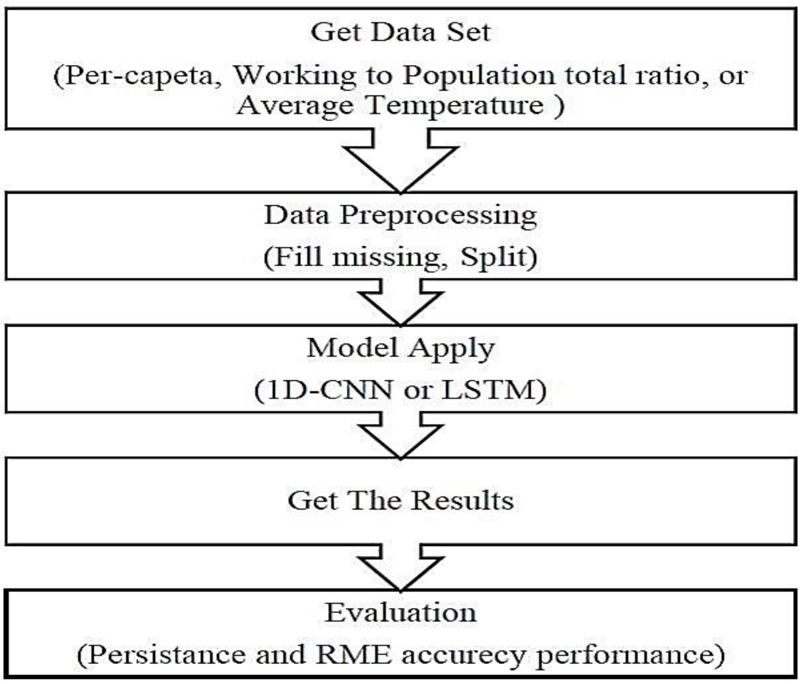

Figure 5. The Flowchart of our proposed approaches.

We use Python programming language, version 3.7, to carry out our experiment. Several Python libraries, including some of the following:

- Pandas Library – For reading Excel data.
- Matplot Library – For visualizing the applied algorithms.
- Numpy Library – For calculations.
- Keras Library- For sequential model.
- Statsmodels Library – For the correlation between our dataset.

Forecasting using deep learning

Deep learning methods consist of standard LSTM, bidirectional LSTM, encoded LSTM, and CNNs. Time series forecasting using LSTM models has been proposed in many different forms According to the estimates of previous research comparing the performances of the multi-step Deep learning methods; we apply single-variable time series multistep prediction using three specific models; 1D-CNN, ConvLSTM, and encoder-decoder LSTM. Regarding the number of filters and the size of the two-dimensional kernel, we can characterize the ConvLSTM as a single layer. (rows, columns). Since we are dealing with a one-dimensional series, the kernel's fixed-row number is always 1. In each of the corresponding models, we use Adam optimizer, and rectifier linear units (Relu)

1. Forecasting using 1D-CNN

In addition to being straightforward and understandable, non-parametric procedures provide outcomes that are comparable to or even superior to those attained by parametric methods without requiring a priori knowledge of data distribution [40]. A basic CNN is made up of three layers: a convolutional layer, a pooling layer, and a fully linked layer. Figure 6 shows a simple feed-forward CNN. The fully connected layer, a traditional neural network output layer, is used to present the findings. The convolutional layer is responsible for extracting the patterns and features, the pooling layer is used to minimize the dimensions. Figure 7 shows a simple convolutional example. Figure 8 shows a pooling example. In order to dramatically minimize the number of inputs in the network, CNN uses local connections between neurons (each neuron is connected to only the nearest neurons of the next layer). Figure 9 shows the artificial neuron that makes up a neural network. A neural network with more than three layers, including the inputs and outputs, is referred to as a "deep learning algorithm". Simply put, a simple neural network is one that just contains two or three layers. Convolutional Neural Networks (CNNs) are thought to be capable of learning from complex data and delivering desired outcomes. Despite being primarily designed for two-dimensional image data, CNNs can model univariate time series forecasting problems.

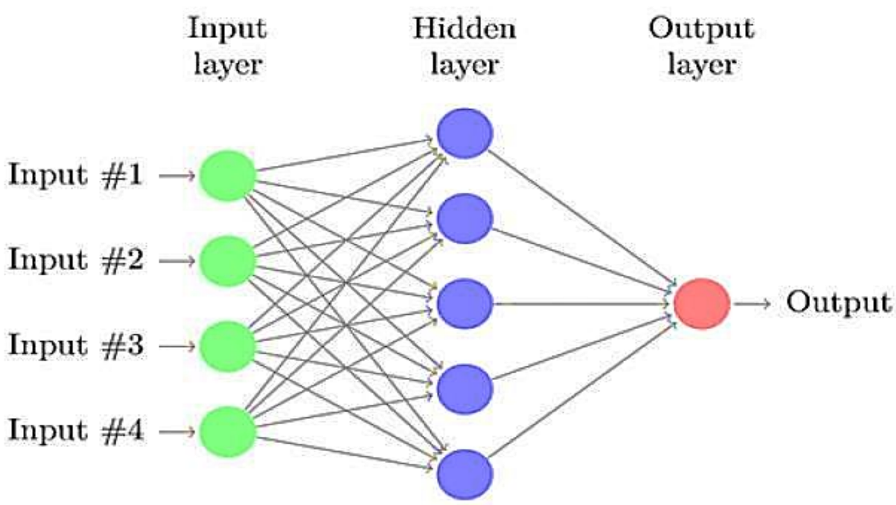

Figure 6 information only ever moves forward in a feed-forward network; it never reverses direction.

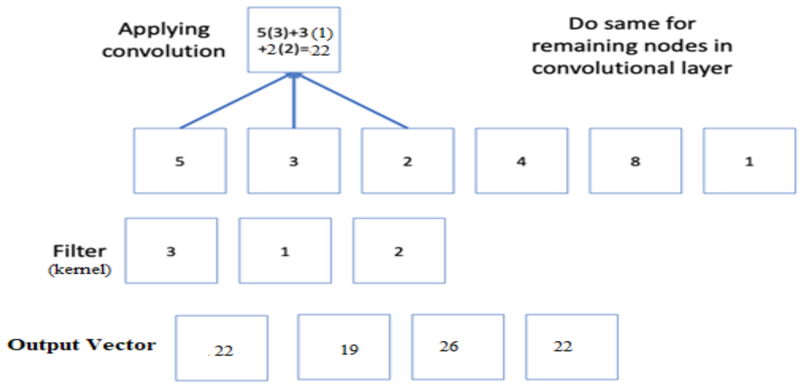

Figure 7 simple example for applying the convolution and the output vector: The filter here is $(\mathbf{kernel}) = (\mathbf{3}, \mathbf{1}, \mathbf{2})$and to apply this to the first three elements $(\mathbf{5}, \mathbf{3}, \mathbf{2})^{T}$ we do $\mathbf{3}(\mathbf{5}) + \mathbf{1}(\mathbf{3}) + \mathbf{2}(\mathbf{2}) = \mathbf{22}$ then apply the same to $(\mathbf{3}, \mathbf{2}, \mathbf{4})^{T}, (\mathbf{2}, \mathbf{4}, \mathbf{8})^{T}, \boldsymbol{and}(\mathbf{4}, \mathbf{8}, \mathbf{1})^{T}$.

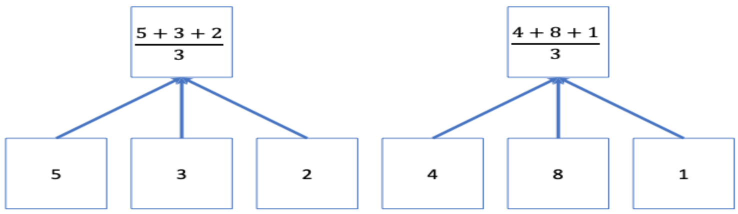

Figure 8 simple average pooling example.

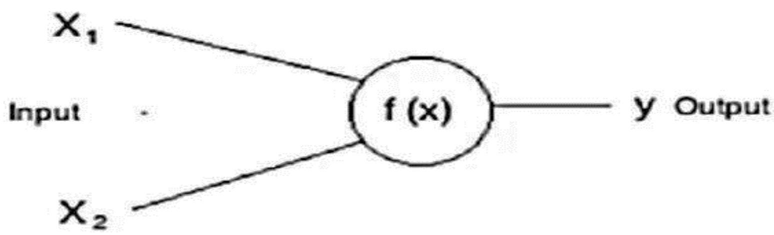

Figure 9 the artificial neurons' inputs are x_1 and x_2, their processing is represented by $f(x)$, and their output is represented by y.

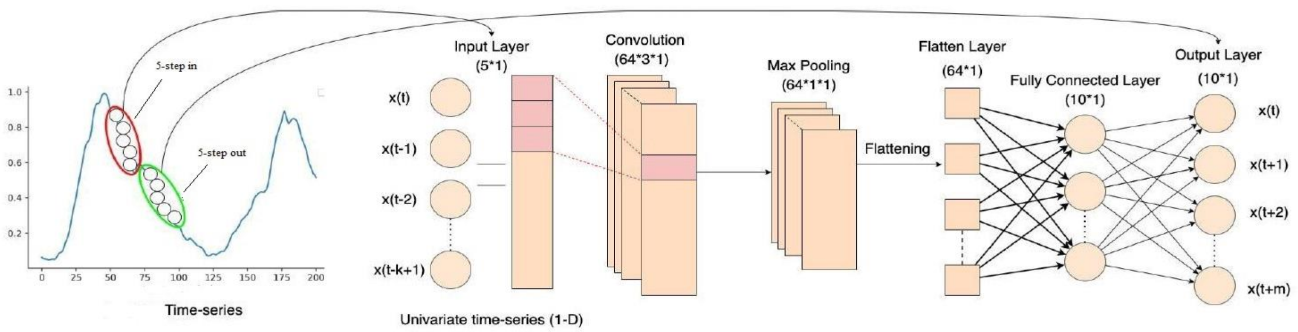

Figure 10 construction of 1D CNN for multi-step ahead time series prediction.

We will implement a CNN sequential model for multi-step time series forecasting utilizing our data, such as the per-capita coffee consumption sequence, as a single variable. Predict the upcoming standard years of this data using a certain number of preceding years' total per-capita coffee consumption. The one-dimensional (1D) subsequence of data that the CNN will read and train to extract features from depends on the number of previous years utilized as input. The data series must be divided into several examples so the model can learn from them .In order to learn an n-step prediction, we can split the sequence into numerous input/output patterns, or samples. This input's/ output's size is denoted “n_steps_in” and “n_steps_out” in the implemented algorithm. If the input's size is five years prior then the “n_steps_in” equals 5 and if the output's size is the next predicted five years, then the “n_steps_out” equals 5. Figure 10 shows a simple example of a construction of 1D CNN for multi-step ahead time series prediction.

2. Forecasting using 1D-ConvLSTM

In a hybrid model with an LSTM backend, a CNN model can be utilized to interpret subsequences of input that are collectively delivered as a sequence to the LSTM model. The input sequences must first be divided into segments that the CNN model can process. For instance, we can first divide our univariate time series data into samples for input and output, using four steps as input and one as output. Then, each sample can be divided into two sub-samples with a total of four-time steps. Each subsequence of two-time steps can be interpreted by the CNN, which can then deliver a time series of interpretations of the subsequences to the LSTM model for processing as input. ConvLSTM, a kind of LSTM related to CNN-LSTM, incorporates convolutional input reading directly into each LSTM unit. The ConvLSTM can be used for univariate time series forecasting even though it was designed for reading two-dimensional spatial-temporal data.

3. Forecasting using encoder-decoder LSTM

The encoder-decoder LSTM approach, compared to the other deep learning methods, also known as sequence-to-sequence (seq2seq), is significantly more effective than just outputting a vector for sequence prediction. The two sub-models that make up an encoder-decoder LSTM are the encoder, which reads input sequences and compresses them to a fixed-length internal representation, and the decoder, which decodes the internal representation and interprets it to predict the output sequence. The multi-step model has the ability to predict results for a single sample. By providing the input, we can forecast the upcoming two or more future time steps.

5. Evaluation

Using a Baseline.

A baseline technique provides a collection of predictions that we can evaluate using the same criteria we would use to evaluate any prediction for our problem, such as prediction accuracy or RMSE. The scores from these algorithms give the crucial point of reference when comparing the outcomes of all other machine learning algorithms to these ones for our challenge. We used the persistence algorithm (naive forecast) as a baseline method in order to forecast what will happen in the following time step (t+1), where the persistence algorithm employs the value from the previous time step (t-1). The first 66% of the observations will be kept for "training," while the final 34% will be used for evaluation. We apply the one-step prediction of the algorithm using the Python programming language, version 3.7. Figure 11 shows the baseline test prediction for the KSA Population total Working Ratio with MSE 1.871. Figure 12 shows the Baseline Test prediction for KSA coffee per-capita consumers with MSE 0.039. Figure 13 shows the Baseline test prediction for the KSA total average temperature with MSE 22.801.

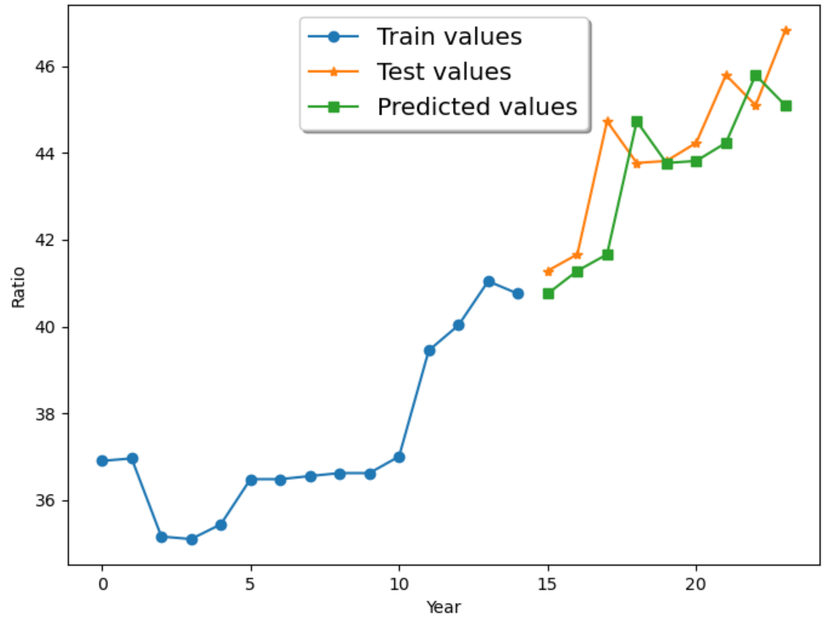

Figure 11 the Persistence Test prediction for KSA Population total Working Ratio with MSE 1.871. The predicted value is the green color curve.

Using a Comparison of RME.

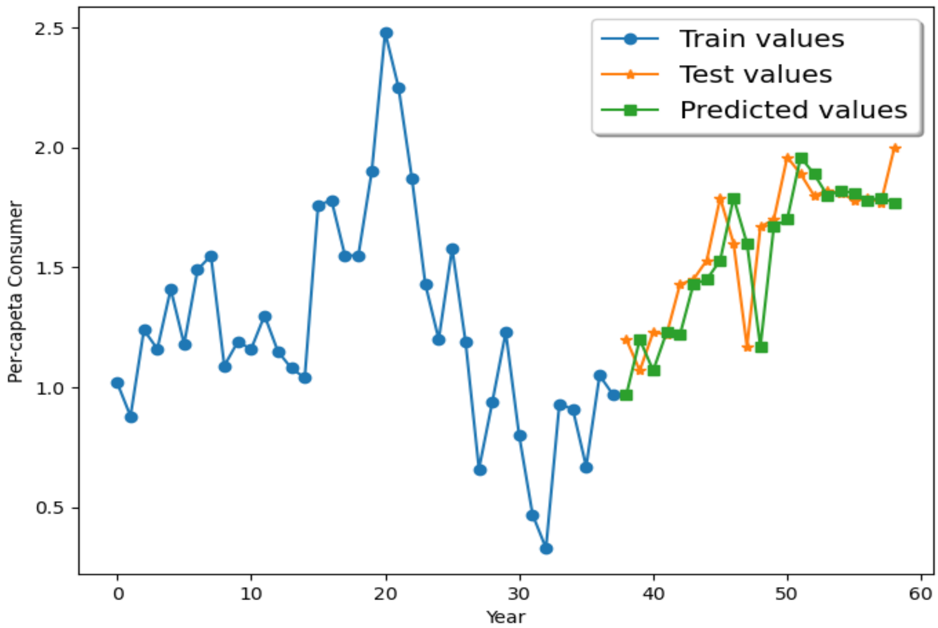

Figure 12 the persistence Baseline Test prediction for KSA coffee per-capita consumers with MSE 0.039. The predicted value is the green color curve.

The average quadratic discrepancies between the expected and actual values are taken into account by MSE. The root of the mean of the squared discrepancies between predictions and actual values is known as the root of the square mean of error or RMSE. With 1D-CNN, we have provided an evaluation of the performance of the Multi-step inputs and multi-step outputs, with 1D-CNN and ED-LSTM, both the input and output components will be formed of several time steps and have the same step count or not. The model anticipates for each sample a vector output that represents multiple time steps of one variable. We run the model a few times and get the average outcome. Tables 4 and 5 show the results of forecasting in terms of accuracy by the two: MSE and RMSE.

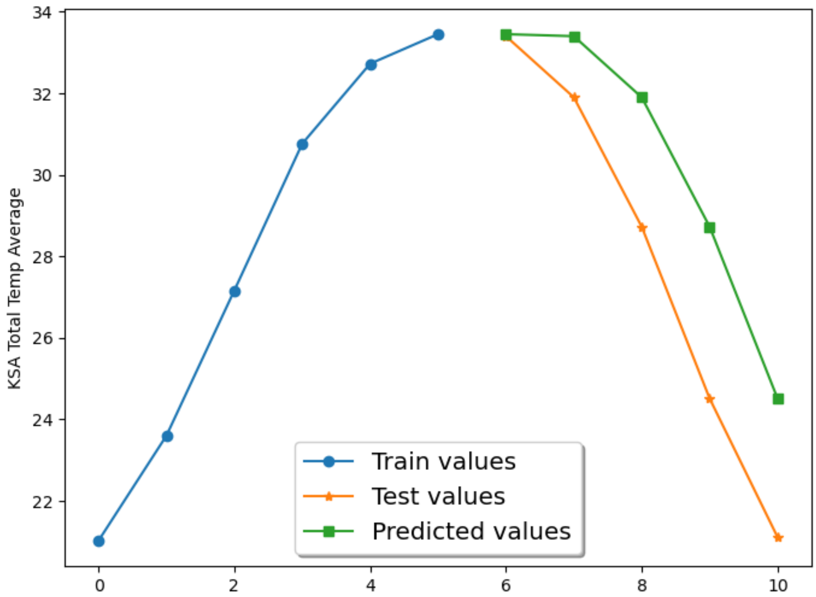

Figure 13 the persistence Baseline Test prediction for the KSA total average temperature with MSE 22.801. The predicted value is the green color curve.

Table 4 the prediction error with 1 output step.

		Persistence algorithm		1D-CNN		1D-ConvLSTM	
Series	n-Steps out	MSE	RMSE	MSE	RMSE	MSE	RMSE
KSA Total average temperature	1	22.801	4.7750	0.00299621	0.0547376	0.69882145	0.69882145
KSA per-capita coffee consumer	1	0.039	0.1974	0.001237929	0.035184221	0.0011655661	0.034140388
Average KSA Population Working Ratio	1	1.871	1.3678	0.02106303	0.145131073	0.003061920	0.0553346

Table 5 the prediction error using 1D-CNN and ED-LSTM with more than 1 output step.

		1D-CNN		ED-LSTM	
Series	n-Steps out	MSE	RMSE	MSE	RMSE
KSA Total average temperature	2	2.59779840	1.611768	0.413325230	0.642903749
	3	0.0318273	0.1784023	1.0677083	1.033299
	4	0.6600746	0.812449	1.63806	1.279868
KSA per-capita coffee consumer	2	0.01279	0.11312	0.03089897	0.175781054
	3	0.0001803	0.013429	0.048729227	0.2207469

	4	0.01867	0.136657	0.124634449	0.353036045
Average KSA Population Working Ratio	2	0.0229064	0.151348	0.135521	0.3681323
	3	0.1053014	0.324501	0.23314117	0.48284695
	4	0.2423837	0.492324	1.2934366	1.1372935

6. Conclusion

Based on the findings, after comparing our suggested multistep one-dimensional neural networks with each other, the study reveals that the multistep one-dimensional convolution neural network 1D-CNN can predict one or more following steps substantially more accurately than convolution Long Short-Term Memory networks or 1D-ConvLSTM, and Encoder-Decoder LSTM or ED-LSTM. The mean square error (MSE) and root mean square error (RMSE) are used to compare the methodologies' accuracy. Tables 4 and 5, listed the forecasting errors of the comparison methods. Future study will apply these technologies to recent advancements in our daily lives, such as paralleling the extraction of expected information with the mining of massive volumes of data.

Acknowledgements

Competing Interests

The author has declared that no competing interests exist.

Appreciation

All appreciation and respect to the Dean of the College for exerting efforts and urging faculty members to increase the submission and publication of scientific research.
All appreciation and respect to the coordinator of scientific research in the college for supporting the faculty members and cooperating with them to facilitate all obstacles.

References

1. Fofana, M., Adounkpe, J., Dotse, S., Bokar, H., Limantol, A., Hounkpe, J., Larbi, I., and Toure, A. (2023) "Flood Forecasting and Warning System: A Survey of Models and Their Applications in West Africa." American Journal of Climate Change, 12, 1-20. doi: 10.4236/ajcc.2023.121001.
2. Al Bassam, A.,Zaidi, F. , Hussein, M.. (2014). Natural hazards in Saudi Arabia. 10.13140/2.1.3599.1046.
3. Nkordeh, N., Popoola, S., Jefia, A., Okeoghene, M., Atayero, A.A.A., Awomoyi, M. and Bobmanuel, I. (2023) The Impact of Satellite Communications on Environmental Hazard Control: Tool for the Realization of African Union Agenda 2063 Aspirations. Int. J. Communications, Network and System Sciences, 16, 191-216. https://doi.org/10.4236/ijcns.2023.168013
4. Akinbobola, A. and Hamisu, S. (2022) "Malaria and Climate Variability in Two Northern Stations of Nigeria." American Journal of Climate Change, 11, 59-78. doi: 10.4236/ajcc.2022.112004.
5. The Top Coffee-Consuming Countries. (2020, August 6). WorldAtlas. https://www.worldatlas.com/articles/top-10-coffee-consuming-nations.html.
6. https://en.wikipedia.org/wiki/List_of_countries_by_coffee_production.n.d.
7. Coffee Consumption Per Capita in Saudi Arabia. (2023, January 12). Coffee Consumption per Capita in Saudi Arabia | Helgi Library. https://www.helgilibrary.com/indicators/coffee-consumption-per-capita/saudi-arabia/
8. https://view.offapps.live.com/op/view.aspx?src=https%3A%2F%2Fwww.stats.gov.sa%2Fsites%2Fdefault%2Ffiles%2FLMS_2022_Q2_AR_1.xlsx&wdOrigin=BROWSELINK.& Methodology Link: https://www.stats.gov.sa/en/node/9787. &Report Link: https://database.stats.gov.sa/home/report/2763.
9. ء. (2023, April 15). الهيئة العامة للإحصاء. https://www.stats.gov.sa/ar & https://www.stats.gov.sa/en/814.
10. Climate & Coffee. (2015, June 19). NOAA Climate.gov. http://www.climate.gov/news-features/climate-and/climate-coffee
11. The Money Trees - CDP. (n.d.). The Money Trees - CDP. https://www.cdp.net/en/research/global-reports/the-money-trees?cid=318635718&adgpid=73254185713&itemid=&targid=aud-841759796550:kwd-

301862320692&mt=b&loc=1028873&ntwk=g&dev=c&dmod=&adp=&gad=1&gclid=Cj0KCQjwlumhBhClARIsABO6p-zijYozgkgAolEkZsM_D7iR4YCEF6d7jUm_kKylyXNreLnnUDqxFrwaApo4EALw_wcB.

12. HTTPS://WWW.CLIMATESTOTRAVEL.COM/CLIMATE/SAUDI-ARABIA. (n.d.). https://HTTPS://WWW.CLIMATESTOTRAVEL.COM/CLIMATE/SAUDI-ARABIA
13. Mitchell, D. C., Knight, C. A., Hockenberry, J., Teplansky, R., & Hartman, T. J. (2014, January). Beverage caffeine intakes in the U.S. Food and Chemical Toxicology, 63, 136–142. https://doi.org/10.1016/j.fct.2013.10.042.
14. Salwa. A. Albar , Almaghrabi MA, Bukhari RA, Alghanmi RH, Althaiban MA, Yaghmour KA. (2021 Jun). Caffeine Sources and Consumption among Saudi Adults Living with Diabetes and Its Potential Effect on HbA1c. Nutrients. 7;13(6):1960. doi: 10.3390/nu13061960. PMID: 34200398; PMCID: PMC8226783.
15. Press release distribution, EDGAR filing, XBRL, regulatory filings | Business Wire. (2023, April 15). Press Release Distribution, EDGAR Filing, XBRL, Regulatory Filings | Business Wire. https://www.businesswire.com
16. Lara, O. D. and Labrador, M. A. (2013). A Survey on Human Activity Recognition using Wearable Sensors," in IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1192-1209, Third Quarter, doi: 10.1109/SURV.2012.110112.00192.
17. Singh, D., Merdivan, E., Psychoula, I., Kropf, J., Hanke, S., Geist, M., & Holzinger, A. (2017). Human Activity Recognition Using Recurrent Neural Networks. Lecture Notes in Computer Science, 267–274. https://doi.org/10.1007/978-3-319-66808-6_18
18. Bengio, Y. (2009). Learning Deep Architectures for AI. Foundations and Trends® in Machine Learning, 2(1), 1–127. https://doi.org/10.1561/2200000006
19. Deng, L. (2014). A tutorial survey of architectures, algorithms, and applications for deep learning. APSIPA Transactions on Signal and Information Processing, 3(1). https://doi.org/10.1017/atsip.2013.9
20. Kiranyaz, S., Ince, T., Gabbouj, M. (2016). Real-time patient-specific ECG classification by 1-D convolutional neural networks", IEEE Trans. Biomed. Eng. 63 664–675, https://doi.org/10.1109/TBME.2015.2468589
21. Kiranyaz, S., Ince, T., Gabbouj, M. (2017). Personalized monitoring and advance warning system for cardiac arrhythmias. Sci. Rep. 7, https://doi.org/ 10.1038/s41598-017-09544-z.
22. Abdeljaber,O. Avci, O. , Kiranyaz S, Gabbouj, M. Inman D.J. (2017). Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. J. Sound Vib. 388 https://doi.org/10.1016/j.jsv.2016.10.043.
23. Ince ,T., Kiranyaz, S., Eren, L., Askar, M., Gabbouj, M. (2016). Real-time motor fault detection by 1-D convolutional neural networks." IEEE Trans. Ind. Electron. 63- 7067–7075, https://doi.org/10.1109/TIE.2016.2582729.
24. Sun, C., Chen, Z., Qin, Y., and Wang, B. (2022). Multi-step Time Series Forecasting Based on InformerXGBoost-GA. Journal of Physics: Conference Series 2333 (2022) 012009. doi:10.1088/1742-6596/2333/1/012009. Chen Sun et al 2022 J. Phys.: Conf. Ser. 2333 012009.
25. Cortez, P., Pereira, P.J., and Mendes, R. (2020) multi-step time series prediction intervals using neuroevolution. Neural Computing and Applications 32(5) 8939-8953.
26. Gao, C., Zhang, N., Li Y, et al. (2022). Self-attention-based time-variant neural networks for multistep time series forecasting Neural Computing and Applications. 34(11) 8737-8754.
27. Campanharo, A., Sirer, M. I., Malmgren, R. D., Ramos, F. M., Amaral, L., and Perc, M. (2011), Duality between time series and networks, Plos One, 6, e23378.
28. Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324.
29. Wang, H.-z., Li, G.-q., Wang, G.-b., Peng, J.-c., Jiang, H., and Liu, Y.-t. (2017). Deep learning-based ensemble approach for probabilistic wind power forecasting. Applied energy, vol. 188, pp. 56–70.
30. Tran, T.V., Yang, B.-S., Tan, A.C.C. (2009). Multi-step ahead direct prediction for the machine condition prognosis using regression trees and neuro-fuzzy systems. Expert Syst. Appl. 36(5), 9378–9387.
31. Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. Computer Vision – ECCV 2014, 818–833. https://doi.org/10.1007/978-3-319-10590-1_53
32. Min, S., Lee, B., & Yoon, S. (2016, July 29). Deep learning in bioinformatics. Briefings in Bioinformatics, bbw068. https://doi.org/10.1093/bib/bbw068

33. Wu, M., Stefanakos, C., Gao, Z., & Haver, S. (2019, September). Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty. Ocean Engineering, 188, 106300. https://doi.org/10.1016/j.oceaneng.2019.106300
34. Zhang, L., Zhou, W. D., Chang, P. C., Yang, J. W., & Li, F. Z. (2013, January). Iterated time series prediction with multiple support vector regression models. Neurocomputing, 99, 411–422. https://doi.org/10.1016/j.neucom.2012.06.030
35. Grigorievskiy, A., Miche, Y., Ventelä, A. M., Séverin, E., & Lendasse, A. (2014, March). Long-term time series prediction using OP-ELM. Neural Networks, 51, 50–56. https://doi.org/10.1016/j.neunet.2013.12.002
36. Ben Taieb, S., Sorjamaa, A., & Bontempi, G. (2010, June). Multiple-output modeling for multi-step-ahead time series forecasting. Neurocomputing, 73(10–12), 1950–1957. https://doi.org/10.1016/j.neucom.2009.11.030
37. Hochreiter, S., & Schmidhuber, J. (1997, November 1). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.
38. https://www.utstat.toronto.edu. (n.d.).
39. VISUAL CROSSING. Weather Data Services | Visual Crossing [online]. https://www.visualcrossing.com/weather/weather-dataservices#/login.
40. Eiglsperger, J., Haselbeck, F., & Grimm, D. G. (2023, June). ForeTiS: A comprehensive time series forecasting framework in Python. Machine Learning with Applications, 12, 100467. https://doi.org/10.1016/j.mlwa.2023.100467

