Journal of Cardiovascular Disease Research

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

A SURVEY ON TIME SERIES PREDICTION USING MACHINE LEARNING

[1]Shivaratri Narasimha Rao ,[2]S.China Ramu

[1]Researcher Scholar, Department of Computer Science and Engineering
Osmania University, Hyderabad, Telangana, India.
getnrao@gmail.com

[2]Professor, Department of Computer Science and Engineering
CBIT(A), Hyderabad-75, Telangana, India.
chinaramu_cse@cbit.ac.in

Abstract

This paper offers an summary of present day literature related to time series classification by implementing using Deep Neural Network (DNN) related methods in initial time series for classifying based on distinct distance procedures similar to Euclidean or active distances based on time warping as the paper initiates by reviewing standard approaches of time series classification which aspires for classifying time series through minimal chronological classifications that are potential for possessing the minimal classification accuracy. Additionally most of the papers emphasis in influencing the essential machine learning procedures as the process explores various machine learning tools that are required to implement various projects of machine learning that describes chronological information combined for predicting each model that highlights recent advancements in hybrid deep learning models that are based on statistical models using with neural network constituents for improving each of the categories of time series data.

Keywords: Deep Neural Networks, Artificial Intelligence, Time Series Data, Prediction

1. Introduction

In order to collect and index the data concept of Time Series arises which is merely based on the aspect of time which is being classified automatically and clusters are being generated are considered as the historical aspect which is being combined with “Deep Neural Networks (DNN) stream” [1]. Although DNN is extensive and thrilling area of “Artificial Intelligence (AI)” in order to perform the research over machines that are elaborated to accomplish training process through series based aspects for solving solve real world issues by implementing various algorithms using various learning algorithms comprises of highest acceptable rate as the DNN and various “Data Mining (DM) techniques tends to explore data from its dataset that comprises of various patterns” [2].

In this paper, we tend to encapsulate various techniques that are popular in regard of time series prediction being implemented in DNN by specifying the profound procedures that are available for performing basic level of forecasting the problems or issues similar to multi prospect forecasting which is very uncertain while being estimate the further analysis that emerges over various trendy novel hybrid models [8]. By combining various models related to quantitative domain specific models using “Deep Learning (DL)” [11] mechanisms for improvising the performance aspects of forecasting with two key approaches that are outlined as the Neural Networks for performing decision support based on prediction models such as interpretability and counterfactual [3].

In this paper the literature review primarily emphases over the previous and present research in the domain of Time series classification by means of DNN which primarily focuses over traditional time series classification [2]. And most of the referred examples of classification comprises of basic practices of time series comprises of areas such as health diagnoses which is traditionally classify the monitoring systems where the data pertaining to patient may lead to initial identification of diseases which can be further prevented by identifying the malfunctioning parts or components or modules similar to identification of machines or their parts may lead to smooth running of a manufacturing unit[4].

2. Literature Review

The Time series data is being measured specifically as the combination of series of distinct observations O_t where the O denotes various forms of observations that are related to time t over the non stop and isolated time series based on the isolated observations [14] that are established that are based on the uninterruptedly logged observations along with their annotations over a specific time span or interval being normalized to be O_0=[0,1].

The sampling rate varies by implementing timeseries prediction the observations attained based on attained results primarily focuses on the literature that is based on the discrete timeseries which comprises of distinct sampling rates or values [5]. The other remarkable feature of Time series is that the attained observations is defined as x_t which is the remarkable singular value which is denoted as "Univariate Time Series (UTS)" [6]. But when the x_t denotes more than one value or a multi dimensional vector denoted as "Multivariate Time Series (MTS)" [7]. As per another research both the UTS and MTS are thoroughly analyzed with comparative results being provided in [8] as the main influential part is the observation comprises of distinct values that are being assumed and need to be defined in R being denoted as R_n which is a combination of set of symbols or alphabet denoted as a data set Ω of finite data [9-12].

The data pertaining to Time series is being applied over numerous contexts for being implemented over the areas such as economic forecasting or to perform stock market analysis or to analyze the impact of inventory being implemented in applications based on their theory aspects to be delt in more detail [13,14].

there are many datasets or repositories that are available for free over internet which comprises of time series data similar to UCR which comprises of sampler data ready to be used for performing analysis of data being classified and archived based on labeled data as UTS as considered to be a benchmark for performing classification [15,16]. As most of the researchers have done research and addressed distinct issues that are based on various time series repository as most of the benchmarks are clearly based on the data source [17,18]. And when dealing with MTS dataset the Australian sign language based data set is being applied and executed for numerous times while performing the time series classification with following analysis of research is being performed [19]:

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

Table 2.1 analysis of existing work

Reference	Proposed Work	methodology	Restrictions
Bhaskaran [19]	Information examination for long period for estimating and booking of hierarchical assets based on time arrangement"	Deterministic model, stochastic model, R dialect instrument	Limited to only few case studies as the prediction is based on very minimal parameters
Petitjean et al. [20]	Arranging and examining the Satellite picture time which is being managed sporadically for the inspected arrangement	Dynamic time warping	Dynamic time warping comprises of very less limitations in terms of templates that are used to train the examples and the levenshtein distance has its own drawbacks which may not support numerical values
Chujai et al. [21]	Examination of model based on time series conjecture power utilization in a family for locating the reasonable estimating period	ARMA models, AIC, and RMSE	Prediction is restricted to only two models for performing forecasting which is performed in half-yearly or yearly basis
Devi et al. [22]	Examination and arrangement of time series for stock based pattern forecast	ARIMA, PMAD, MAPE, % blunder and precision	The accuracy of the error percentage is around 16.24% which is being considered to be very impressive
Smith et al. [23]	Gathering of patent data based on Innovation for determining the models	HWES and ARIMA	After exploring various associations that are grouped based on various validation factors related to patents data"
Boubacar et al. [24]	The probability technique implemented over sustainable power with hot spots being used with greater insight over administration of microgrid framework	Wavelet deterioration and fake neural systems	Reduction of resource percentage with ANN consists of difficult model for interpreting the lesser performance on small datasets for computationally intensive to train
Laptev et al. [25]	Investigation of event-based forecasting using neural networks for Uber	Estimating the vulnerability and heterogeneous gauging	Based on length of time series for correlating among those time series are low then it is not feasible

3. Methods in Machine Learning

Over past years an enormous number of ML algorithms were being presented as only very few of them which has the potential to solve the problem are being replaced by another ones [20] similar to unsupervised learning and reinforcement learning and one more is the supervised learning being displayed as one of the following.

3.1 Supervised learning: is clearly based one few learning techniques which comprises of distinct set of input variables being represented through training data are actually pre labeled as per the target data being considered as an input variable based on which the mapping functions are required for p of the producing the results based on needed outputs by implementing the parameter modifying techniques being implemented with certain level of suitable accuracy based on teaching data acquired [21].

3.2 Unsupervised learning: In this learning technique of algorithm there exists only training data instead of the outcome of data considered as input and the one which is not earlier labeled as it makes use of certain types of recognized classifiers implemented over existing patterns or cluster in a input datasets [22].

3.3 Reinforcement learning: is implemented when the algorithm tends to train the machine based on the requirements such as updating of reward or feedback the signals are generated and further the machine will be trained based on itself for identifying the most remunerating the actions based on the experience [23].

4. Deep Learning for Time Series Classification

The "Time series forecasting" is implemented over predictive models based on the future values of a target data being denoted by y_i, t aimed at a given entity i over the time t as the each of the entity denotes a logical alignment of chronological information which is similar to measuring the calculations attained from individual sources of climate condition stations being performed in the area of climatology based on the generated vital signs that differentiates patients from medicine and further the whole process can be easily observed by performing the one-step-ahead forecasting using the form [24]:

$$y_i, t+1 = f(y_i, t-k{:}t,\ x_i, t-k{:}t, s_i) \tag{4.1}$$

Where the generated result of equation 4.1 will yield the predicted model based on the values calculated from the right hand equation being denoted with various distinct observations that are performed towards obtaining the target and exogenous inputs individually executed using the process of look-back window denoted by statistical based metadate with k, s_i for representing a sensor location or an object location and similarly $f()$ function is implemented for performing the prediction function for attaining the knowledge form the model. The primary focus will be implied over the univariate forecasting with 1-D targets over similar components implemented and extended using the multivariate models and without loss of generality by omitting the index i over subsequent sections [25].

The DNNs are trained to perform prediction over relationships and their factors by implementing time series over non-linear layers for generating the intermediate features by performing encoding of appropriate historical information with latent variable which is produced by forecasting is denoted by z_t

$$f(y_t - k{:}t,\ x_t - k{:}t, s) = g_{desc}(z_t) \tag{4.2}$$

$$z_t = g_{enc}(y_t - k{:}t,\ x_t - k{:}t, s) \tag{4.3}$$

In equation 4.2 and 4.3 the $g_{enc}()$ and $g_{desc}()$ functions are implemented for performing the encoding and decoding process in functions respectively and further the calculation of recall with the subscript i is obtainged from Equation (4.1) and can be easily removed for simplifying the notation (e.g. $y{:}t$ substituted by y_t) and these encoders and decoders comprises of essential building blocks of various DL architectures based on the selection process of determining network categories of associations being learnt by the model proposed.

4.1 Convolutional Neural Networks

"Convolutional neural networks (CNN)" is a extraction tool for acquiring the local relationships over a traditional image dataset which is invariant across spatial dimensions adapted over the Time Series Datasets as most of the researchers will make use of numerous layers of causal convolution filters intended to guarantee the past information being used for performing forecasting over the intermediate feature at a hidden layer l with each of the causal convolutional filter as:

$$h_t^{l+1} = A((W * h)(l,t)) \tag{4.4}$$

$$(W * h)(l,t) = \sum_{T=0}^{k} W(l,T)\, h_{t-T}^{l} \tag{4.5}$$

Where L.H.S of equation 4.4 is the intermediate state over the layer l with the over time t and * being denoted while performing the convolutional operation denoted as W(l, T) which belongs to the constant amount of filter weight over a particular layer l with further function who symbolizes the activation function which is almost similar to sigmoid function in any of the provided or used architectures comprising of non-linear execution of data and the output of CNN which is denoted by $z_t = h_t^l$

By specifying the 1-D case we can identify that the Equation (4.5) comprises of a robust likeliness to "Finite Impulse Response (FIR)" which are also considered as the filters being implemented in digital signal processing with two key associations over the chronological association learnt by CNNs with in lining the spatial invariance that are assumed over the regular CNNs. With the association of temporal CNNs we tend to assume that the association of time invariant based filter weights set over each of the time stamps. In addition to this CNNs are capable of implementing the algorithm using the defined inputs over the defined lookback window for implementing the forecast over receptive fields of size k which is required to be tuned carefully for ensuring the relevancy of historical information using the linear activation function [26].

4.2 Recurrent Neural Networks

The "Recurrent neural networks (RNNs)" is traditionally being used in performing modelling in sequential order [5] where the process comprises of strong the results over a variation of NLP tasks [6] being interpreted naturally over the time series data which comprises of sequences of data in the form of inputs and targets s most of the RNNs are based on the developed architectures for performing the forecasting of core applications [12] with an internal memory state required to act as the compact summary related to the previous information in a recursive manner which is frequently updated with attained novel observations over each of the step.

$$z_t = v(z_t - 1, y_t, x_t, s) \tag{4.6}$$

In the equation 4.6 the L.H.S which is z_t related to the anticipated hidden internal state of RNN where the function v() comprises of the learnt memory being denoted and updated with Elman RNN over distinct variants is evaluated as:

$$y_{t+1} = \gamma_y (W_y z_t + b_y) \tag{4.7}$$

$$z_t = \gamma_y \left(W_{z_1} z_{t-1} + W_{z_2} y_t + W_{z_3} x_{t+ W_{z_4}} s + b_z\right) \tag{4.8}$$

Where W and b are denoted as linear weights which are based on the network with functions $\gamma_y()$, $\gamma_z()$ respectively are the network activation functions which are not required in RNNs which explicitly specify the lookback window in terms of CNN in the view point of signal implementation which is the foremost recurring layer based on Equation 4.8 which resembles the various non-linear version of "Infinite Impulse Response (IIR)" filters.

5. Early Time Series Forecasting

Based on the observations generating a classification decision is considered to be a major issue over the timeseries data which may lead to lose the classification accuracy with distinct interpretations that the previous classification problem is delivered to tackle the computational performance for imparting the classification. And this paper will improvise in terms of both time and resource complexity using 1-Nearest Neighbor by implementing dynamic time warping distance measure for generating the nearest "centroid" classifiers which are faster and not that much accurate when compared with nearest neighbor algorithms implemented over the multivariate time series data sets which acts as a benchmarking technique for comparing various approaches specified [12].

Initial works being implemented in this area is the primary classification being defined as the over time with possible series length [16]. The literals are considered to be the simple indicators which illustrates statistically whether the time series data is going up or down based on various pre-defined intervals that are combined with AdaBoost comprising of ensemble classifiers that are generated to be implemented over multivariate time series through variable lengths implemented on ensemble classifier with distinct literals to attain the classification response as output [17]. It is a fact that an ensemble or collaborative classifier is a linear combination of literals with unknown values generated based on the classification decisions made over the partial sample data in time series for classifying the data [18]. When datasets are considered such as: CBF, Control, Trace and Auslan (Gloves) the generated classification error rate (minimum) is 0.43% which is further reduced by 75% of its original length (maximum) is 85% is further reduced by 65% of its original length along with the boosted interval literals lead to achieve the boast average error rates of 13% for the 65% time series length case and a minimal of 1.52% for the 85% case [19].

The More probabilistic framework is proposed by name "RelClass" which comprises of quadratic discriminant usage to support vector machines that performs classification as the major aspect is reliability of decision taken by performing classification with a certain degree of confidence over the incomplete data is considerably sufficient over the complete data in the framework that comprises of classification performed with a threshold value [20].

6. Algorithms of Machine learning for Time series prediction

There exists many substantial amount of algorithms being implemented over the machine learning and are implemented based on their learning procedures are:

6.1. Regression algorithms

In Regression algorithms the prediction is performed over the model based on the relationship modeling between distinct variables for measuring the error rate in an iterative manner using the regression technique [5]. The variable may be an attribute such as price or a specific temperature being favored with the regression algorithms are:

ISSN: 0975-3583, 0976-2833 VOL 12, ISSUE 03, 2021

- “Linear Regression algorithm” [1]
- “Ordinary Least Squares Regression” [2]
- “Multivariate Adaptive Regression Splines” [3]
- “Logistic Regression” [4]
- “Locally Estimated Scatter plot Smoothing” [5]
- “Stepwise Regression” [6]

6.2. Illustration based learning algorithms

The illustration based algorithms are needed for performing the decision problem which is an issue for illustrating the training data for constructing the database which compares the test data for perform prediction as the illustration based learning technique is better known as lazy learner and some of the instance based algorithms are:

- “Learning Vector Quantization” [7]
- “Self-Organizing Map” [8]
- “k-Nearest Neighbor” [9]
- “Locally Weighted Learning” [10]

6.3. The Algorithms implement based on decision Tree

Algorithms that comprise of various decision trees that are implemented primarily for resolving classification problem by splitting the attributes in deuce or more sets being categorized based on their values and generating tree with variant nodes and variant branches and the algorithms that makes use of decision tree are:

- “Iterative Dichotomized” [11]
- “M5” [12]
- “Chi squared Automatic Interaction Detection” [12]
- “C5.0 and C4.5” [13]
- “Decision Stump” [14]
- “Classification and Regression Tree” [15]
- “Conditional Decision Trees” [16]

6.4. Baysian algorithms

The field of ML is a multidisciplinary area being researched in Computer Science similar to Statistics and algorithm where the statistics accomplishes and enumerates the vagueness represented by “Bayesian algorithms” that are merely based on “probability theory and Bayes theorem and the algorithms” that makes use are:

- “Bayesian Belief Network (BBN)” [17]
- “Multinomial Naive Bayes Bayesian Network (BN)” [18]
- “Averaged One-Dependence Estimators (AODE)” [19]
- “Gaussian Naive Bayes” [20]
- “Naive Bayes” [21]

6.5. Algorithms being implemented using ANN

The prototypes that are based on ANN are similarly established on the configuration of “biological neuron” by implementing the supervised learning which comprises of weighted artificial neurons for establishing interconnections between various units in parallel distributed processing networks and the algorithms that makes use of ANN are:

- “Radial Basis Function Network (RBFN)” [22]
- “Back-Propagation” [23]
- “Perceptron” [24]
- “Hopfield Network” [25]

6.6. Algorithms using Dimensionality Reduction

When there exists large volumes of data leads to space as a concern leads to Dimensionality reduction based on the statistical significance for reducing the numerous dimensions that are being outlined to be the nominal item which tends to remove dissimilar and unessential data based on the reduced processing cost and the algorithms that makes use of dimensionality reduction are:

- "Partial Least Squares Regression" [10]
- "Multidimensional Scaling" [13]
- "Principal Component Analysis" [14]
- "Flexible Discriminant Analysis" [16]
- "Mixture Discriminant Analysis" [19]
- "Sammon Mapping" [2]
- "Projection Pursuit" [12]
- "Linear Discriminant Analysis" [11]
- "Principal Component Regression" [18]
- "Quadratic Discriminant Analysis" [22]

7. Tools Used in Machine Learning

Swift and Rapid are the tools that delivers machine learning interface to be used through a programming language that makes a machine to learn by providing the best practices for development process with distinct platforms are capable of executing a module and the machine learning tools are:

- "Python SciPy subparts such as scikit-learn (Panda)" [6]
- "R Platform" [7]
- "WEKA Machine Learning Workbench" [8]

And the ML tools comprises of distinct libraries with complete project capabilities and some of the most often used libraries are:

- "JSAT in Java" [2]
- "scikit-learn in Python" [5]
- "Accord Framework in .NET" [6]

8. Conclusion

In this paper we have thoroughly reviewed and analyzed distinct machine learning methods and algorithms which have been defining with distinct kinds of ML techniques, tools, algorithms and methodology along with various applications of ML that are required for processing the data implemented in this paper. We have also identified the research work done by referring almost 25 papers in this area which comprises of distinct machine learning algorithms being implemented in this decade with discrete conventional methods which have been outperformed in the previous models.

References

1. Teng Xiuyi1,Gong Yuxia1. "Research on Application of Machine Learning in Data Mining". IOP Conf. Series: Materials Science and Engineering, 2018.
2. M. Praveena,V. Jaiganesh, "Literature Review on Supervised Machine Learning Algorithms and Boosting Process". International Journal of Computer Applications, ISSN No. 0975 – 8887, vol. 169, 2017.
3. Mohammed Ali Shaik 2019 A Survey on Text Classification methods through Machine Learning Methods *International Journal of Control and Automation IJCA* Vol (12) Issue-6 Pp 390-396
4. Kajaree Das, Rabi Narayan Behera. "A Survey on Machine learning: Concept, Algorithms and Applications", International Journal of Innovative Research in Computer and communication Engineering . vol. 5, 2017.
5. Mohammed Ali Shaik Dhanraj Verma P Praveen K Ranganath and Bonthala Prabhanjan Yadav 2020 RNN based prediction of spatiotemporal data mining *IOP Conf Ser: Mater Sci Eng* Vol (981) 022027 https://doiorg/101088/1757-899X/981/2/022027
6. Rahim Barzegar, Asghar Asghari Moghaddam, Jan Adamowski, Bogdan Ozga-Zielinski,"Multi-step water quality forecasting using a boosting ensemble multi-wavelet extreme learning machine model", Springer, ISBN 00477-017-1394, 2017.

7. Alexis Fouilloy , Cyril Voyant , Gilles Notton, Fabrice Motte ,Christophe Paoli, Marie-Laure Nivet,, Emmanuel Guillot, JeanLaurent Duchaud ," Solar irradiation prediction with machine learning: Forecasting models selection method depending on weather variability", Energy 165, 2018.
8. Spyros Makridakis, Evangelos Spiliotis, Vassilios Assimakopoulos, " Statistical and Machine Learning forecasting methods: Concerns and ways forward", PLoS ONE 13,2018.
9. Mohammed Ali Shaik T Sampath Kumar P Praveen R Vijayaprakash 2019 Research on Multi-Agent Experiment in Clustering International Journal of Recent Technology and Engineering IJRTE Vol (8) Issue-1S4 June 2019 Pp 1126-1129
10. Yu Zhang, Yu Wang, Guoxu Zhou, Jing Jin, Bei Wang, Xingyu Wang, Andrzej Cichocki, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces", Expert Systems With Applications 96, pp 302–310, 2018.
11. Mohammed Ali Shaik 2020 Time Series Forecasting using Vector quantization International Journal of Advanced Science and Technology IJAST Vol (29) Issue-4 Pp 169-175
12. Teng Xiuyi1, Gong Yuxia1, Research on Application of Machine Learning in Data Mining, IOP Conf. Series: Materials Science and Engineering 392 (2018)
13. Mudelsee M. Trend analysis of climate time series: A review of methods. Earth-Science Reviews. 2019;190:310 – 322.
14. Mohammed Ali Shaik and Dhanraj Verma 2020 Deep learning time series to forecast COVID-19 active cases in INDIA: a comparative study IOP Conf Ser: Mater Sci Eng Vol (981) 022041 https://doiorg/101088/1757-899X/981/2/022041
15. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nature Medicine. 2019 Jan;25(1):44–56.
16. Böse JH, Flunkert V, Gasthaus J, Januschowski T, Lange D, Salinas D, et al. Probabilistic Demand Forecasting at Scale. Proc VLDB Endow. 2017 Aug;10(12):1694–1705.
17. Devlin J, Chang MW, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers); 2019. p. 4171–4186.
18. Mohammed Ali Shaik and Dhanraj Verma 2020 Enhanced ANN training model to smooth and time series forecast IOP Conf Ser: Mater Sci Eng Vol (981) 022038 https://doiorg/101088/1757-899X/981/2/022038
19. Bhaskaran S (2012) Time series data analysis for long term forecasting and scheduling of organizational resources-few cases. Int J Comput Appl 41(12):4–9
20. Petitjean F, Inglada J, Ganc¸arski P (2012) Satellite image time series analysis under time warping. IEEE Trans Geosci Remote Sens 50:3081–3095. https://doi.org/10.1109/TGRS.2011.2179050
21. Chujai P, Kerdprasop N, Kerdprasop K (2013) Time series analysis of household electric consumption with ARIMA and ARMA models. In: Proceedings of the international multi conference of engineers and computer scientists, IMECS 2013, Hong Kong, 13–15 Mar 2013, vol 1, pp 295–300
22. Devi BU, Sundar D, Alli P (2013) An effective time series analysis for stock trend prediction using ARIMA model for nifty midcap-50. Int J Data Min Knowl Manag Process 3:65–78. https://doi.org/10.5121/ijdkp.2013.3106
23. Smith M, Agrawal R (2015) A comparison of time series model forecasting methods on patent groups. In: CEUR workshop proceedings, pp 167–173
24. Doucoure B, Agbossou K, Cardenas A (2016) Time series prediction using artificial wavelet neural network and multi-resolution analysis: application to wind speed data. Renew Energy 92:202–211. https://doi.org/10.1016/j.renene.2016.02.003
25. Laptev N, Yosinski J, Li LE, Smyl S (2017) Time-series extreme event forecasting with neural networks at Uber. In: International conference on machine learning, Sydney, Australia, 11 Aug 2017, vol 34, pp 1–5