Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

DROUGHT STAGE PREDICTION FROM REMOTE SENSING BASED VEGETATION AND WATER INDEXES USING MACHINE LEARNING

MUHAMMAD OWAIS RAZA
Masters, University of Sufism and Modern Sciences Bhitshah, Pakistan.
E-mail: owais.leghari@hotmail.com

ZULFIQAR ALI BHATTI
Associate Professor, Chemical Engineering Department, Mehran University of Engineering and Technology Jamshoro, Pakistan. E-mail: zulfiqar.bhatti@faculty.muet.edu.pk

MOHSIN MEMON
Associate Professor, Software Engineering Department, Mehran University of Engineering and Technology Jamshoro, Pakistan. E-mail: Mohsin.memon@faculty.muet.edu.pk

SANIA BHATTI
Professor, Software Engineering Department, Mehran University of Engineering and Technology Jamshoro, Pakistan. E-mail: sania.bhatti@faculty.muet.edu.pk

NAZIA PATHAN
Ph.D. Scholar, Queen Margaret University Edinburgh. E-mail: naziapathan09@gmail.com

Abstract

Drought is one of the most uncertain and perilous natural disasters which is caused due to the rapid climate changes that is dampening and worsening with each passing day in terms of its intensity and frequency. In this context, drought modelling is of immense importance, keeping in view its highly negative impact on the affected community globally. Drought has four stages, which are, drought, normal (conditions), pre-drought and after-drought. In this study, drought stages are predicted employing a two-part strategy. The first part of strategy is the forecasting of the remote sensing-based water and vegetation indexes (Enhanced Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI), and Normalized Difference Water Index (NDWI)), while in the second part, drought stages are predicted using EVI, NDVI and NDWI. Unfortunately, Pakistan is one of the most drought-prone regions with a densely populated desert called, “Thar” in Tharpakar, Sindh, Pakistan. Therefore, in this study, the dataset is collected from one of the most drought effected district (Tharpakar) of Pakistan. In this study, ARIMA (Auto-regressive Integrated Moving Average) is used the forecasting of water and vegetation indexes, and, Multiclass Decision Forest (MDF), Multiclass Decision Jungle (MDJ), Multiclass Logistic Regression (MLR) and Multiclass Neural Network (MNN) are employed for the classification of drought stages. Based on the experiments and various assessments performed in this research, the RMSE for EVI ranges between 0.035 and 0.088, for NDVI, it falls between 0.034 and 0.075, and for NDWI, it ranges between 0.032 and 0.075. The best performing algorithm for drought stage prediction is Multi-class Decision Forest with an accuracy of 97.35%

Index Terms: Drought prediction, machine learning, satellite images, drought indexes, EVI, NDVI, and NDWI

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

1. INTRODUCTION

With the rapid rise in Earth temperature, the rate of water circulation has increased due to which natural disasters like drought have become more frequent and occurring at a very fast pace. Consequently, drought occur due to the lack of availability of water causing serious and critical problems in the agriculture production. Ultimately, the food insecurity occurs which results in the worst form of negative socio-economic effects [1] on the people of that region. Drought is considered as one of the most lethal yet uncontrollable disasters, but it is undermined and considered negligibly in our society. Recently, droughts have been occurring frequently in both developed and under-developed 1 nation such as drought in California, USA from 2012 to 2014 [2] and in East Africa from 2010 to 2011 [3]. Pakistan has also faced some drastic droughts in the past such as drought in Balochistan province from 1997 to 2003 and drought in the Sindh province in 2014. In Sindh province, more than 2.5 million people are affected by drought either directly or indirectly every year and 22.7% population of Tharpakar district is at a level of acute malnutrition since 2018-2019 [4]. Drought leaves a huge impact on Pakistan's economy since Pakistan is an agricultural country and agriculture comprises of 21% of its total GDP [5].

Drought prediction is one of the biggest challenges for meteorologists, climatologists, and environmentalists. Without a proper forecast, prediction and planning, drought can cause huge damage to human life and capital resources. Prediction of drought is very vital for decision-making, as, it can alert and caution stakeholders to take some proactive strategies to handle the disastrous situation of drought effectively and prudently. In Pakistan, the calamity control agencies such as PDMA (Provincial Disaster Management Authority) and NDMA (National Disaster Management Authority) perform forecasting and prediction of drought using various statistical and hybrid models for better decision-making in a bid to avert the chances of growing droughts and other related calamities.

Remote sensing of the drought indexes is one of the methods that have been widely used for the forecasting and prediction of drought and the latest addition to this approach is the intense use of machine learning [6], [7], [8] for accurate predictions. The indexes such as Standardized Precipitation Index (SPI) and Standardized Precipitation-Evapotranspiration Index (SPEI) obtained via remote sensing techniques are generally used to predict and forecast drought via machine learning algorithms.

In this study, EVI, NDVI and NDWI are forecasted for drought assessment using ARIMA, while the algorithms used for the prediction of drought stages are chosen based on the stationary nature of indexes. The classification model developed for this study is named as Model-I. It employees MDF, MDJ, MLR and MNN algorithms on EVI (Enhanced Vegetation), NDVI (Normalized Difference Vegetation Index), and NDWI (Normalized Difference Water Index) for the classification of four drought stages (Drought, Normal (conditions), Pre-drought and After-drought). MLR provides a solid prediction using classical ML approach, while MDJ and MDF are perfect example of ensemble over a

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

classical decision tree. MNN shows the modelling capability of neural network which is suitable for this task. The main contributions of this study are listed below:

Curating and Annotation of the vegetation and water based remote sensing indexes (EVI, NDVI, and NDWI) dataset of Tharpakar District for forecasting and classification of drought stages (Normal, Drought, Pre-drought, and Post-drought). As far as our knowledge no such dataset exists.

Creation and evaluation of ARIMA model on vegetation and water indexes (EVI, NDVI, and NDWI) for drought modelling on Tharpakar District.

Creation and Evaluation of Model-I for drought stages classification with EVI, NDVI, and NDWI using MLR, MDJ, MDF and MNN algorithms. To the best of our knowledge this only study that performs drought stage classification using remote sensing indexes.

2. RELATED WORK

Drought prediction and forecasting are one of the biggest challenges of the 21st century because of the climatic and environmental issues and the drought’s adverse implications over the humans and other living beings [9]. The efficacy of machine learning algorithms has given researchers a productive way to predict the drought effectively, timely, and accurately [10]. Researchers have employed a variety of methods of data handling to find co-relations and used those correlations for the forecasting of the drought using SPI. SPI is also used in conjunction with SPEI by [6] to forecast indexes to predict the drought. They have used ARIMA and LSTM and suggested that ARIMA is considerably better and viable at providing good results as well as computationally less complex than LSTM. It has been evaluated that the biggest hurdle in creating a machine learning model for drought is to obtain the relevant data. In [7] researchers developed a machine learning framework to detect and predict drought using vegetation indexes data which was extracted from Landsat Dataset.

To achieve better results, researchers have used a wide variety of machine learning algorithms; as in [8] multi-layer perceptron neural network (MLPNN), multiple linear regression (MLR), and co-active neuro-fuzzy inference system (CANFIS) were used for drought prediction using EDI (Effective Drought Index). Some other techniques like wavelet transforms were also coupled and correlated with machine learning algorithms to increase the prediction accuracy [11] In (Landsat EVI), Researchers have performed drought prediction for Pakistan with SPEI using machine learning during Rabi and Kharif seasons. In this regard, the research for drought prediction in Pakistan is at its initial phase. Authors in [12] have created a drought index called Combined Terrestrial Evapotranspiration Index (CTEI) with the help of five machine learning algorithms SVM, Decision Trees, Bagged Trees, Boosted Trees, and Gaussian process regression on Ganga River Basin. These algorithms used a combination of hydro-meteorological inputs and satellite data. Through the observed results, the best-performing model was found to be SVM. Development of drought modelling framework is also an essential viewpoint for prediction of drought as in [13] researchers have designed and evaluated drought modelling framework using committee extreme learning machine, committee multiple linear regression and committee particle swarm optimization-adaptive neuro-fuzzy inference system. These models were used to forecast SPI (Standard Precipitation Index) every month. The chosen test centres for the study are Multan, Islamabad, and Dera Ismail Khan. The heuristic approaches can also be used for the drought use case as in, [14] MLPNN, Multiple Linear Regression, and Co-active Neuro-Fuzzy Inference System for hydrological drought prediction in Uttarakhand State, India on upper Ramganga River basin. The best performing model in the study was MLPNN.

To understand the sufficient time scale for drought monitoring, researchers in [13] used the Boruta algorithm and random forest to perform drought monitoring with the help of the Standardized Precipitation Temperature Index (SPTI) from fifty-two meteorological stations which are located at different places in Pakistan. Fog and Cloud computing in the field of Artificial Intelligence (AI) is a powerful tool, as the researchers in [15] proposed a Fog Cloud-based prediction framework for drought prediction using NN and Genetic algorithms. The preliminary results of this work employing supervised machine learning algorithms on satellite images are presented in [16].

After having a careful review of the literature, it is evident that a wide variety of parameters are used for drought prediction using remote sensing indexes. However, most of the studies focus on a single index for the forecasting and map the results of that index to drought. To deal with the weaknesses of previously proposed techniques of drought modelling for drought stage classification, in this research dataset of EVI, NDVI and NDWI is curated in such a way that, both, forecasting of individual index and classification of drought stages can be performed. ARIMA is used for the forecasting and various ML and DL methods are used for the classification of drought stages. The area under the study is Tharpakar district Pakistan.

3. STUDY AREA

The area selected for this study is Tharpakar, Sindh, Pakistan. Tharparkar is located between 69° 3 35 E and 71° 7 47 E longitudes, and between 24° 9 35 N and 25° 43 6 N latitudes. The population of Tharparkar according to the 2017 census is 165, 966, 1. Tharparkar is a deserted region with a low number of fertile lands. It has been the epicentre of many droughts in the past [17] which is the main reason for choosing this as the area of study. There are seven talukas in the Tharparkar district namely Chachro, Dali, Diplo, Islamkot, Kaloi, Mithi, and Nagarparkar.

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

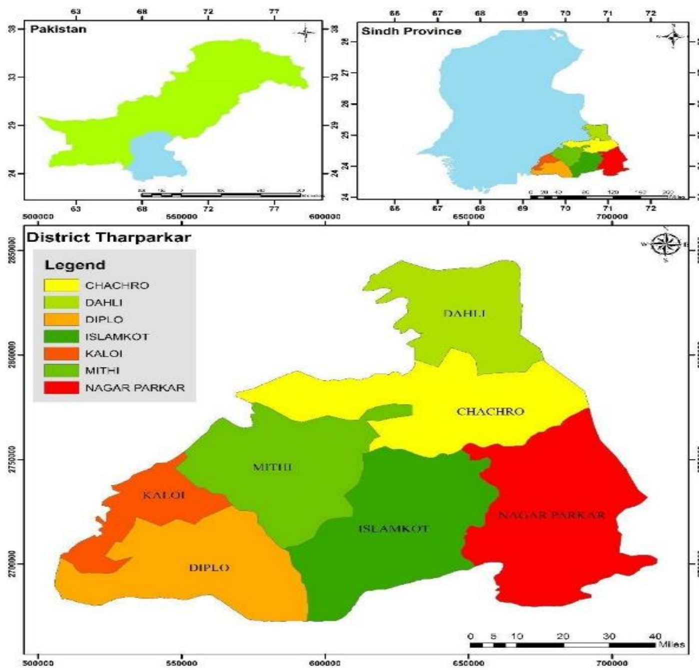


Fig 1: Map of Tharparkar District Showing Seven Talukas

4. WATER & VEGETATION INDEXES

In this study, EVI, NDVI, NDWI, are used for drought modelling for all seven talukas of Tharparkar district. Each of the indexes is discussed comprehensively in this section.

4.1 Enhanced Vegetation Index (EVI)

The Enhanced Vegetation Index is the measure of greenness, which is used to quantify vegetation greenness. It works well with most of the environments, while with canopy background, more noise is generated for dense vegetation. EVI is generated from NIR (Near Infrared), Red, and blue bands, and, is calculated as the ratio of R and NIR. Its value ranges from -1.0 to 1.0. Eq (1) represents EVI, here NIR represents Near-Infrared Band and R shows Red Band [18]

$$EVI = \frac{2.5 * ((NIR - R)}{(NIR + 6.0 * R - 7.5 * B + 1.0))} \qquad (1)$$

4.2 Normalized Difference Vegetation Index (NDVI)

Normalized Difference Vegetation Index is derived from Thematic Mapper (TM). It is used to measure the greenness and density of vegetation. It falls in between Red and Near-Infrared and is calculated as a ratio between R and NIR. Eq (2) represents NDVI in which NIR is near-infrared and R is the red band [19]

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \qquad (2)$$

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

4.3 Normalized Difference Water Index (NDWI)

The Normalized Difference Water Index is used as the water content at the canopy level and is one of the important indexes to be taken into consideration for drought prediction. It is derived from NIR and Short-Wave Infrared (SWIR), and it is calculated as the ratio of NIR and SWIR. Eq (3) represents NDWI, here NIR is Near Infrared and SWIR is Shortwave Infrared Band [20]

$$NDWI = \frac{(NIR - SWIR)}{(NIR + SWIR)} \quad (3)$$

5. METHODOLOGY AND IMPLEMENTATION

In this study, the forecasting of indexes for drought modelling and the prediction of drought stages is performed. Figure 2 shows the detailed methodology and implementation steps for this study and in this section each block of methodology is comprehensively discussed.

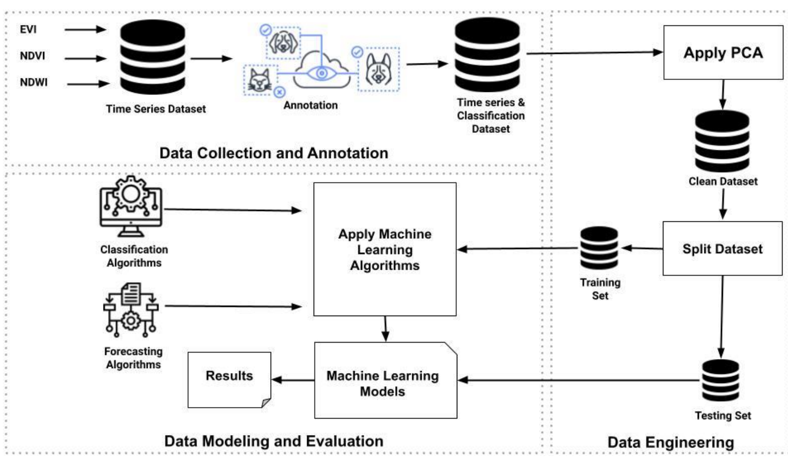


Fig 2: Implementation methodology diagram

5.1 Data Collection and annotation

The first part in the implementation of a machine learning model is the availability of the data. For the study, vegetation, and water indexes EVI, NDVI and NDWI of all seven talukas are collected for the prediction of drought stages. Data of all the indexes are collected from 1987 to 2020 using Landsat collections. In this research, Landsat 5 and Landsat 7 are used [21] where Landsat 7 is exploited for the data from 1987 to 2012 and Landsat 5 is used to fill gaps in the data between 1999 and 2012 as well as for providing the data from 2012 to 2020. The resolution that Landsat uses is 30 meters of the Earth's surface and takes the data including multi-spectral and thermal data. The data is being collected with an interval of 32 days. There are 12 observations for all seven 8 talukas in one year; overall the collected data is for 33 years. Table.1 shows the mean and median in EVI, NDVI, and NDWI values for all seven talukas in the dataset. The data in the dataset

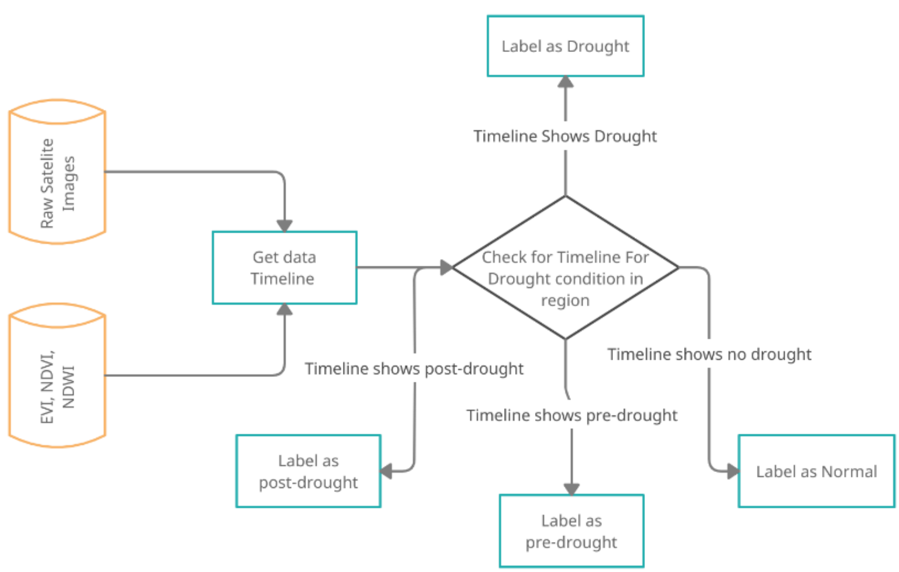
Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

is timestamped and it is further annotated into four labels, namely, Normal, Drought, Pre-drought, and Post-drought. 'Normal' represents no drought. 'Drought' means that the drought is at the peak, and it is not diminishing to go towards normal, whereas, 'pre-drought' shows conditions before the drought and that the drought is on its way. 'Post-drought' means that the drought has ended, and things are going back to normal. Each instance of indexes must be labelled as one of above-mentioned classes. Fig.3 represents the flow for labelling of each instance in the dataset. The timestamp in the dataset allows for the forecasting indexes individually and drought stage label is used for the classification of drought stages.

Table 1: Mean, Median for EVI, NDVI, and NDWI of Tharparkar District

District	EVI Mean	NDVI Mean	NDWI Mean	EVI Median	NDVI Median	NDWI Median
Chachro	0.099	0.092	-0.054	0.085	0.082	-0.071
Dali	0.082	0.096	-0.057	0.072	0.083	-0.074
Diplo	0.082	0.107	-0.031	0.071	0.099	-0.05
Islamkot	0.11	0.089	-0.061	0.103	0.079	-0.081
Kaloi	0.116	0.115	-0.049	0.11	0.101	-0.103
Mithi	0.096	0.098	-0.059	0.087	0.091	-0.078
Nagarparkar	0.102	0.131	-0.043	0.86	0.12	-0.061

Fig 3: Flow of Annotation the Process



ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

5.2 Data Engineering

Once the dataset is collected and properly annotated, the next step is to clean this dataset from any missing values. PCA is used to deal with the missing values in this research [22].

5.2.1 Principal Component Analysis (PCA)

PCA is a dimensionality reduction technique that reduces a dataset with many features to a small one while containing most of the data from the original dataset. In this research, missing values are replaced via PCA which uses a linear model to estimate the approximation of data by analysing the correlations between the columns. It also has some additional advantages of other imputation techniques, such as, Replace by Mean, Replace by Median, and Multiple Imputation using Chained Equations (MICE) as it does not require the application of a predictor for all the columns, rather, it provides covariance for the full dataset [23].

For this study, the dataset is prepared for the forecasting and classification in two different ways. For forecasting, the training dataset consists of EVI, NDVI and NDWI from 1987 to 2017 and for testing, the dataset is taken from 2018 to 2020. For drought stage classification, the dataset is splited into 80%, 20% subsets for training and testing respectively. The selection of ratios is based on the results from different split ratios. 80-20 was found to be the most efficient splitting ratio in this research study.

5.3 Data Modelling and Evaluation

Since we have the data ready to be fed into the algorithm, two types of modelling and evaluation approaches are used here. One for forecasting called as the Forecasting and Evaluation Scenario, 10 and, second for drought stage classification mentioned here as the Classification and Evaluation Scenario.

5.3.1 Forecasting and Evaluation Scenario

In this section, the modelling approaches and evaluation metrics used in the forecasting of EVI, NDVI and NDWI are discussed along with the details of ARIMA forecasting model and relevant evaluation metrics.

Auto-regressive Integrated Moving Average (ARIMA): An auto-regressive integrated moving average, or ARIMA, is a statistical analysis model that uses the time series data to either better understand the dataset or to predict the future trends. A statistical model is auto regressive when it predicts the future values based on the past values. For example, an ARIMA model might seek to predict a stock’s future prices based on the past performance or forecast a company’s earnings based on the past periods. After applying ARIMA, the next step is to perform evaluation for the forecasting. Following are the evaluation parameter for the forecasting used in this study.

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

Augmented Dickey Fuller Test ADF: The Augmented Dickey-Fuller (ADF) test is a type of statistical test called a unit root test [24]. The ADF works with the higher-order auto-regressive processes which are represented by Eq(4) where yt is the data under study.

$$\Delta y_t \;= \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \ldots + \delta_p \Delta y_{t-p} + \varepsilon_t \qquad (4)$$

α is a constant

β is a coefficient on trend t

γ Represents the coefficient of first lag of y

ε_t Represents random stationary disturbance term at t

δ_p Represents the coefficient of the first difference of y at lag t-1?

β1 is a constant; β2 is a slope coefficient on time trend *t*

H_0: $\Delta y_t \neq$ stationary

H_1: $\Delta y_t =$ stationary

Akaike information criterion (AIC): Hirotugu Akaike, prominent Japanese statistician, developed the (Akaike information criterion) AIC. AIC is a metric for comparing the time series models that are used in statistics. The AIC, however, is not really a test statistic, hence it cannot ensure a model's quality when compared with the other models. If all the models under consideration fit poorly to a set of data, the AIC will merely suggest the model that fits the data somewhat better than the others. AIC is well expressed in Eq (5).

$$AIC = \frac{-2L}{N} + \frac{2K}{N} \qquad (5)$$

Bayesian information criterion (BIC): The (Bayesian information criterion) BIC, also known as the Schwarz Criterion, is a statistical metric which is used to analyse the time series models. A statistician, Gideon Schwarz devised BIC and it is closely related to the AIC. When we improve the model's quality of fit by adding k parameters, the contrast between BIC and AIC become clearer. In this case, the BIC penalizes such a parameter rise more harshly than AIC. Eq (6) represents BIC.

$$BIC = \frac{-2L}{n} + \frac{-K \ln N}{n} \qquad (6)$$

Similarly, in Eq (6), L is log likelihood function, N is number of observations, K is the number of interpreters and regressors.

Hannan-Quinn information criterion (HQIC): The Hannan-Quinn information criterion (HQIC), developed by Edward Hannan and Barry Quinn, employs yet another cost function that is extremely close to the BIC, as stated in Eq (7).

$$HQIC = \frac{-2L}{n} + \frac{-K \ln \ln N}{n} \qquad (7)$$

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

Root Mean Squared Error (RMSE): The root mean squared error (RMSE) is the square root of the mean of all squared errors. The RMSE error measure is popular, and it is frequently considered as a useful all-purpose error measure for the numerical predictions. Eq (8) depicts RMSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \qquad (8)$$

In this equation, O_i is the actual prediction and S_i represents the predicted values. The main purpose of RMSE is to compare models to acquire the one with least errors.

5.3.2 Classification and Evaluation Scenario

In this section, the modelling approaches and evaluation metrics used in the forecasting of EVI, NDVI and NDWI are discussed along with the details of ARIMA forecasting model and relevant evaluation metrics.

A detailed discussion of machine learning algorithms used for classification in this study are presented in this section. The algorithms are measured based on the classification evaluation metrics such as accuracy, precision and recall.

Multiclass Logistic Regression (MLR): Logistic Regression is a classification algorithm which assumes that the data relies on logistic distribution. The standard form of logistic regression is binomial, however, when dealing with multiclass, it assumes each class as binary. In this research, there are four classes and standard logistic regression works by considering each class as if it exists or not, like, if it is Normal/Drought/Pre-drought or Not Normal/No drought/No pre-drought [25].

Multiclass Decision Forest (MDF): Multiclass decision forest is an algorithm that uses multifarious learning techniques for classification. Multiclass decision forest works by creating multiple decision trees, and, after the creation of decision trees, the best output class is selected by voting. In this algorithm, voting is the class of aggregation in which a non-normalized label's histogram is represented as an output of every decision tree. Further, these histograms are summed and normalized to get probabilities for each label. The trees having the best prediction yield more weightage in the final ensemble model. This model has been employed during our research because, it can work well with non-linear decision boundaries and integrated features [26].

Multiclass Decision Jungle (MDJ): Multiclass Decision Jungle is a classification algorithm that is the extension of the decision forest. Decision jungle is an ensemble of DAGs (decision-directed acyclic graphs), and, like decision forest, it works well with integrated features and non-linear decision boundaries which suits our research. Moreover, it provides better generalization than decision forests [27].

Multiclass Neural Network (MNN): The Multiclass Neural Network is used as a classification algorithm in this research which is considered as a supervised machine

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

learning task. This type of neural network is good for performing complex tasks where the label/output is non-binary. The use of four labels during the experiments makes it suitable for this study [28].

After modelling for classification, evaluation metrics such as accuracy, precision and recall are determined.

Accuracy: The accuracy of a classifier is a performance metric that measures how well it fits across all classifications. It is helpful because all the classes are treated equal by this metric. It is calculated as the proportion of correct predictions to total predictions as shown in the Eq (9).

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \tag{9}$$

Here $T_p, T_n, F_p, and\ F_n$represent True positive, true negative, false positive and false negative respectively

Precision: The precision is calculated as the ratio of correctly identified positive samples to both true and false positive samples classified. Precision is a statistic that measures accuracy of a classifier in predicting whether a sample is positive. Eq (10) represents it.

$$Percision = \frac{TP}{TP + FP} \tag{10}$$

Recall: The recall is calculated as the ratio of correctly classified positive samples to all accessible positive cases. The recall parameter determines how successfully the model detects positive samples. The higher the recall, the more positive samples are detected. Eq (11) represents it

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

6. EXPERIMENTAL RESULTS AND DISCUSSION

After the experiments, the obtained results fall into two categories. The results which are obtained from forecasting model are termed as forecasting results, whereas, the other set of results obtained from the classification model are called as classification results. The experiments were performed on a Core I7 3 GHz processor with an 8 GB random access memory (RAM). The Python programming language was used to implement the machine learning algorithms and other processing tasks.

6.1 Forecasting Results

In this section, the forecasting results are represented. These results are drawn from the test set which consists of data from year 2018 to 2020 (2020 included). The following section discusses each evaluation scheme in detail

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

6.1.1 ADF Results:

In this test, the null hypothesis represents that the series is not stationary, and the alternate hypothesis suggests that the series is stationary [29]. Thus, if the null hypothesis is rejected, it means that the series is stationary and, if it fails to be rejected, it means that the series is not stationary. To test the null hypothesis, the p-value is checked whether it is below a given threshold value which is 0.05. Table 2 shows ADF results for all seven talukas related to 3 indexes which are EVI, NDVI, and NDWI. The decision of acceptance or rejection of the null hypothesis is based on the p-value shown in the table 2. If the p-value is below 0.05, it means that the null hypothesis is rejected, and the series is stationary, otherwise, non-stationary. Further 1%, 5%, and 10% critical values for EVI, NDVI, and NDWI are also calculated. The critical values are compared with the test statistics and if they are less than the critical values, the hypothesis is stated as accepted with 99%, 95%, and 90% confidence. Differentiation transforms the series to make it stationary by stabilizing the means and removing the changes, thus, reducing the seasonality and the trends. It is apparent from Table. 2 that the significance values for all the talukas are negative for EVI, NDVI, and NDWI which indicate that we are 99% confident about the rejection of the null hypothesis

Table 2: ADF Values

Index	Taluka	Test Statistic	p-value	Critical Value (1%)	Critical Value (5%)	Critical Value (10%)
EVI	Chachro	-1.16E+01	2.44E-21	-3.45E+00	-2.87E+00	-2.57E+00
	Dali	-4.52286	0.000179	-3.447142	-2.868941	-2.570713
	Diplo	-4.345409	0.00037	-3.447057	-3.447057	-2.868904
	Islamkot	-2.8781	0.047955	-3.447142	-2.868941	-2.570713
	Kaloi	-2.003149	0.285194	-3.447099	-2.868923	-2.570703
	Mithi	-5.33667	0.000005	-3.447229	-2.86898	-2.570733
	Nangarparkar	-4.702579	0.000083	-3.447142	-2.868941	-2.570713
NDVI	Chachro	-2.159366	0.221322	-3.447057	-2.868904	-2.570693
	Dali	-4.335608	0.000385	-3.447186	-2.86896	-2.570723
	Diplo	-2.516514	0.111508	-3.447229	-2.86898	-2.570733
	Islamkot	-4.55196	0.000158	-3.447229	-2.86898	-2.570733
	Kaloi	-1.569875	0.498706	-3.447057	-2.868904	-2.570693
	Mithi	-6.48E+00	1.28E-08	-3.45E+00	-2.87E+00	-2.57E+00
	Nangarparkar	-4.952163	0.000028	-3.447229	-2.86898	-2.570733
NDWI	Chachro	-4.091205	0.001001	-3.447273	-2.868999	-2.570743
	Dali	-3.900214	0.002033	-3.447057	-2.868904	-2.570693
	Diplo	-2.29909	0.172262	-3.447229	-2.86898	-2.570733
	Islamkot	-3.621266	0.005367	-3.447317	-2.869018	-2.570754
	Kaloi	-1.05E+01	1.20E-18	-3.45E+00	-2.87E+00	-2.57E+00
	Mithi	-2.998952	0.034992	-3.447099	-2.868923	-2.570703
	Nangarparkar	-5.153606	0.000011	-3.447186	-2.86896	-2.570723

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

6.1.2 ARIMA Results

ARIMA model is applied after converting all non-stationary series to stationary series [17] for forecasting EVI, NDVI, and NDWI relating to 7 talukas of Tharparkar. The evaluation parameters selected for the model are Log Likelihood, Akaike's Information Criterion (AIC), Bayesian Information Criterion (BIC) and Hannan–Quinn information criterion (HQIC) [30] Log-Likelihood represents the maximum chances of any estimation made by the model considering the logs of the previous or the last value. This is used to compare the number of models; the higher the log-likelihood means better the model [31] The log-likelihood for our model is -22943.28. The log-likelihood is negative when the likelihood is a number less than one. However, it is not used as a standalone metric in this study, and this metric could be used for comparison with any future model. The next parameter for the model evaluation is AIC which is used to evaluate the robustness of the model by taking the highest likelihood of all the parameters. The best parameter is selected with the lowest AIC values and its value is 45898.561. When we have a model with a complicated parameter; BIC is a better evaluation parameter, and it works like AIC but also considers several rows of the dataset. In our model the BIC value is 45937.555. The HQIC is a rare criterion, but very helpful in feature selection; its values show how efficiently a model is performing with the current features. The value for HQIC is 45912.24. In this section RMSE of ARIMA for all the indexes and talukas is discussed.

Table 3: ARIMA Model Summary

	Log-Likelihood	AIC	BIC	HQIC
Arima Model	-22943.28	45898.561	45937.555	45912.24

Fig.4 shows the RMSE values for all the three indexes, EVI, NDVI, and NDWI, relating to seven talukas of the Tharparkar district. It is evident from fig. 4 that the EVI for Dali taluka has the least error chances and NDWI for Kaloi has the highest RMSE. The RMSE for EVI ranges between 0.035 and 0.088, RMSE for NDVI falls between 0.034 and 0.075, besides RMSE for NDWI ranges between 0.032 and 0.075.

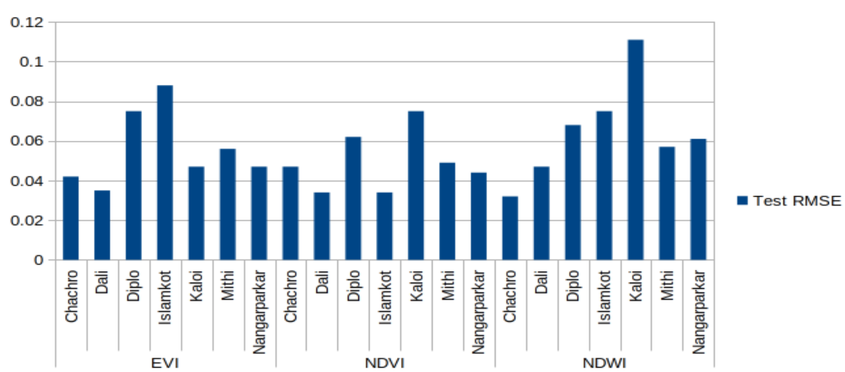


Fig 4: Bar Chart for RMSE Values

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

6.1.3 Monthly forecasting of EVI NDVI and NDWI

Figure 5 shows the monthly prediction and original values of EVI, NDVI, and NDWI. The red colour line shows the distribution of original values while the blue colour line shows the distribution of the values for predicted values of indexes. It is observed from these graphs that the predicted and original values are close to same in the months from November to August but deviate between August and November. This is due to the unpredictable monsoon season which is throwing prediction off the original values.

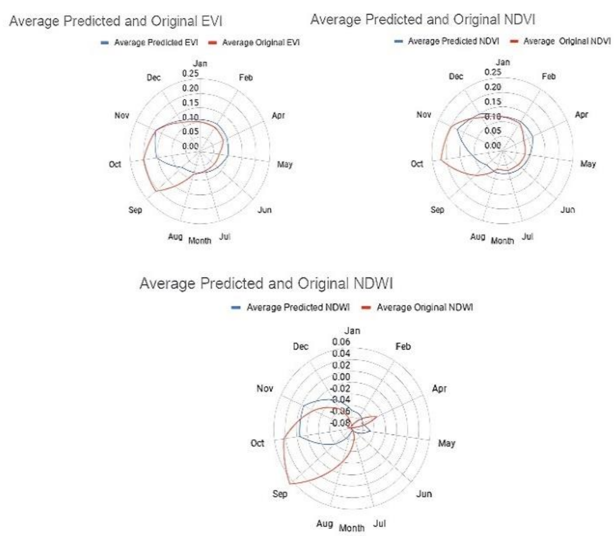


Fig 5: Polar charts for predicted and original EVI, NDVI and NDWI

6.1.4 Regional forecasting of EVI NDVI and NDWI

Figure 6 shows the original and predicted values of EVI, NDVI, and NDWI for the year 2020. The year selection is based on the varying amount of rainfalls and other environmental parameters. The year with highest uncertainty is chosen for prediction due to which year 2020 was chosen. It was the most difficult year in terms of prediction, therefore, to test the model on such data will truly evaluate the model not only on a single region, but all 7 talukas. On the map in figure 6, the high values are represented by darker colours. The original and predicted values are compared in the map, and it is observed that, similar the colour means closer the values. The range of values in the map can be seen in the legend present on each map. In general, legends on each graph represents the value range of that index for a given colour. The lesser the difference between the original and predicted map, the better the model. From figure 6, it is evident that the EVI forecasting for Chachro and Mithi talukas are not very accurate as observed by the difference of colour, but all other forecasts seem to be very close to original values. For NDVI, difference in colour value is observed for Mithi taluka, while the deviation from original values forecasting for all other talukas is up to the mark. For NDWI, Diplo and Kaloi talukas are showing difference in the original and forecasting values, while, for the rest of the talukas, the forecasting values are very close to the original values.

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

6.2.1 Classification Results on 80-20% split:

Table. 4 shows the evaluation parameters of the algorithms, MDF, MDJ, MLR and MNN used in the Model-I. The best performing algorithms for Model-I are MDJ and MDF with the accuracy 21 of 97% and the worse performing algorithm is MLR with the accuracy of 47%. Although these parameters show the overall model performance, it is also important to understand that how models are performing on individual classes. To show the model performance on individual classes, a confusion matrix is plotted as shown in figure 7. Based on the confusion matrix, pre-drought is the most difficult class to predict because it overlaps both normal and drought conditions. Models are exceptionally good at predicting drought and normal classes because these classes don't represent a change. Timely detection of changes in the classification problem is a difficult task, therefore, the models struggle with pre-drought and after-drought classes.

Table 4: Accuracy, Precision and Recall for All the models

Multiclass Decision Forest	Accuracy	0.973592
	Precision	0.973592
	Recall	0.973592
Multiclass Decision Jungle	Accuracy	0.97007
	Precision	0.97007
	Recall	0.97007
Multiclass logistic Regression	Accuracy	0.471831
	Precision	0.471831
	Recall	0.471831
Multiclass Neural Network	Accuracy	0.71831
	Precision	0.71831
	Recall	0.71831

Xi'an Shiyou Daxue Xuebao (Ziran Kexue Ban)/
Journal of Xi'an Shiyou University, Natural Sciences Edition
ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

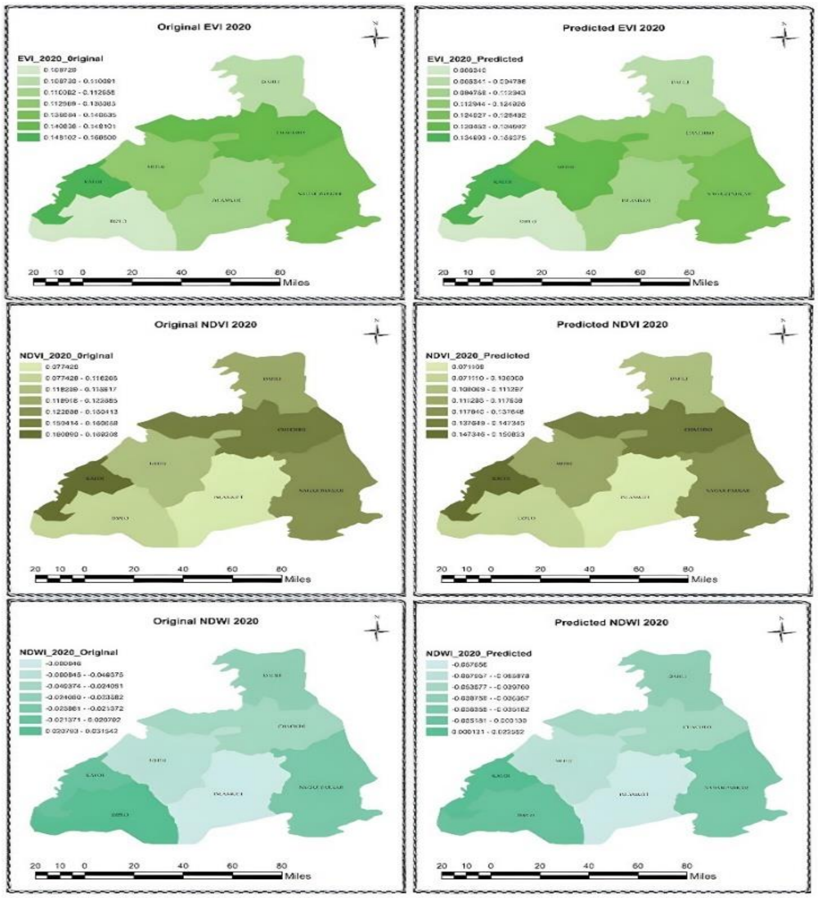


Fig 6: Maps for predicted and original EVI, NDVI and NDWI

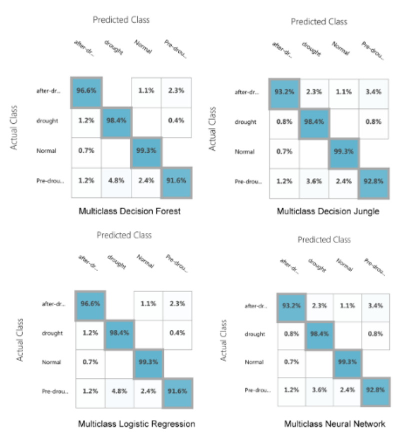


Fig 7: Confusion matrix of various machine learning algorithms

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

7. CONCLUSION & FUTURE WORK

Drought is one of the most perilous natural disasters, as it affects humans and nature equally. Due to the rapid climate changes, the chances of drought conditions have increased manifold. Predicting the drought is one of the biggest challenges even for environmentalists and meteorologists. The conditions of drought on the Earth can be understood by looking at water levels and vegetation in any region. In the light of this fact, one of the most common ways to predict or forecast drought is with the help of indexes such as water, vegetation, and drought indexes. But these techniques are unable to tackle problem of change detection i.e., pre-drought stage (from normal to drought). The area selected for the study is the Tharparkar because it is one of the most densely populated deserts and is very much prone to drought. In this study, a machine learning model is developed called as Model-I. It uses EVI, NDVI, and NDWI for the prediction of drought stages, however, obtaining these indexes in real-time itself is a challenge. Therefore, a forecasting model using ARIMA is developed that can make forecasting of EVI, NDVI and NDWI for several months in advance with a small margin of error. It can also be seen from the literature that many other researchers have also used one or a combination of indexes to perform drought prediction. However, to the best of our knowledge, the combination of indexes i.e., NDVI, NDWI, and EVI used to perform drought stage classification in this research is novel and completely authentic. The complete experiment and research study is based on the data gathered from all seven talukas of Tharpakar district for 33 years starting from 1987 to 2020. A variety of machine learning algorithms were applied on the Model-I and performance of these algorithms were measured. The algorithms selected for this study are Multiclass Logistic Regression, Multiclass Decision Forest, Multiclass Decision Jungle and Multiclass Neural Network. For the proposed Model-I, the best performing model is Multiclass Decision Forest with an accuracy of 97.35%. The data used in this study is unable to make predictions in real-time from day to day since it was recorded with an interval of 32 days. Therefore, in future, this technique will be applied to the data which is available in days and real-time forecasting and drought stage classification will be made available via interactive map.

8. ACKNOWLEDGEMENTS

This work is the extension of the research conducted to fulfill the requirement of a master’s degree. The authors would like to thank the Department of Software engineering, Mehran University of Engineering and Technology for the provision of resources required for the completion of this work. Special thanks to Mr. Mudasir for proofreading the article and Dr. Naeem Mahoto for his comments on the initial draft.

References

1) 1 M. Dai, S. Huang, Q. Huang, et al., “Assessing agricultural drought risk and its dynamic evolution characteristics,” Agricultural Water Management 231, 106003 (2020).

2) A. P. Williams, R. Seager, J. T. Abatzoglou, et al., “Contribution of anthropogenic warming to california drought during 2012–2014,” Geophysical Research Letters 42(16), 6819–6828 (2015).

ISSN: 1673-064X
E-Publication: Online Open Access
Vol: 66 Issue 03 | 2023
DOI 10.17605/OSF.IO/GBD2H

3) V. Ongoma, T. Guirong, B. Ogwang, et al., “Diagnosis of seasonal rainfall variability over east africa: a case study of 2010-2011 drought over kenya,” Pakistan Journal of Meteorology Vol 11(22) (2015).

4) “Who situation report - pakistan: Drought in balochistan and sindh (6 february 2019) - pakistan,” (2019).

5) A. Azam and M. Shafique, “Agriculture in pakistan and its impact on economy,” A Review. Inter. J. Adv. Sci. Technol 103, 47–60 (2017).

6) S. Poornima and M. Pushpalatha, “Drought prediction based on spi and spei with varying timescales using lstm recurrent neural network,” Soft Computing 23(18), 8399–8412 (2019).

7) S. Y. J. Prasetyo, K. D. Hartomo, M. C. Paseleng, et al., “The machine learning to detect drought risk in central java using landsat 8 oli remote sensing images,” in 2019 5th International Conference on Science and Technology (ICST), 1, 1–6, IEEE (2019).

8) A. Malik and A. Kumar, “Meteorological drought prediction using heuristic approaches based on effective drought index: a case study in uttarakhand,” Arabian Journal of Geosciences 13(6), 1–17 (2020).

9) Z. Hao, V. P. Singh, and Y. Xia, “Seasonal drought prediction: advances, challenges, and future prospects,” Reviews of Geophysics 56(1), 108–141 (2018).

10) B. Zahraie, M. Nasseri, and F. Nematizadeh, “Exploring spatiotemporal meteorological correlations for basin scale meteorological drought forecasting using data mining methods,” Arabian Journal of Geosciences 10(19), 1–15 (2017).

11) A. Belayneh, J. Adamowski, B. Khalil, et al., “Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction,” Atmospheric research 172, 37–47 (2016).

12) Elbeltagi, N. Kumari, J. K. Dharpure, et al., “Prediction of combined terrestrial evapotranspiration index (ctei) over large river basin based on machine learning approaches,” Water 13(4), 547 (2021).

13) M. Ali, R. C. Deo, N. J. Downs, et al., “Multi-stage committee based extreme learning machine model incorporating the influence of climate parameters and seasonality on drought forecasting,” Computers and electronics in agriculture 152, 149–165 (2018).

14) A. Malik, A. Kumar, and R. P. Singh, “Application of heuristic approaches for prediction of hydrological drought using multi-scalar streamflow drought index,” Water Resources Management 33(11), 3985–4006 (2019).

15) Kaur and S. K. Sood, “Cloud-fog based framework for drought prediction and forecasting using artificial neural network and genetic algorithm,” Journal of Experimental & Theoretical Artificial Intelligence 32(2), 273–289 (2020).

16) O. Raza, M. Memon, S. Bhatti, et al., “Drought prediction with raw satellite imagery and ensemble supervised machine learning,” Review of Environment and Earth Sciences 8(1), 1–7 (2021)

17) 17 S. Siami-Namini, N. Tavakoli, and A. S. Namin, “A comparison of arima and lstm in forecasting time series,” in 2018 17th IEEE international conference on machine learning and applications (ICMLA), 1394–1401, IEEE (2018).

18) “Landsat Enhanced Vegetation Index | U.S. Geological Survey.”

19) “Landsat Normalized Difference Vegetation Index | U.S. Geological Survey.”

20) “Normalized Difference Water Index: NDWI Formula and Calculations,” (2021).

21) A. V. Egorov, D. P. Roy, H. K. Zhang, et al., “Landsat 4, 5 and 7 (1982 to 2017) analysis ready data (ard) observation coverage over the conterminous united states and implications for terrestrial monitoring,” Remote Sensing 11(4), 447 (2019).

22) L. Balzano, Y. Chi, and Y. M. Lu, “Streaming pca and subspace tracking: The missing data case,” Proceedings of the IEEE 106(8), 1293–1310 (2018).

23) Y.-Y. Choi, H. Shon, Y.-J. Byon, et al., “Enhanced application of principal component analysis in machine learning for imputation of missing traffic data,” Applied Sciences 9(10), 2149 (2019).

24) K. P. Bertelsen, “Comparing the supremum augmented dickey fuller and log periodic power law frameworks for identifying bubbles,” Available at SSRN 3392208 (2019).

25) G. Gasso, “Logistic regression,” (2019).

26) 26 S. Rajagopal, K. S. Hareesha, and P. P. Kundapur, “Performance analysis of binary and multiclass models using azure machine learning.,” International Journal of Electrical & Computer Engineering (2088-8708) 10(1) (2020).

27) Y.-J. NAM and W.-J. SHIN, “A study on comparison of lung cancer prediction using ensemble machine learning,” Korea Journal of Artificial Intelligence 7(2), 19–24 (2019).

28) V. Pliuhin, M. Pan, V. Yesina, et al., “Using azure maching learning cloud technology for electric machines optimization,” in 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), 55–58, IEEE (2018).

29) K. Ajewole, S. Adejuwon, and V. Jemilohun, “Test for stationarity on inflation rates in nigeria using augmented dickey fuller test and phillips-persons test,” J. Math 16, 11–14 (2020).

30) G. Jain and B. Mallick, “A study of time series models arima and ets (january 13, 2017),” Available at SSRN 2898968.

31) E. S. Karakoyun and A. Cibikdiken, “Comparison of arima time series model and lstm deep learning algorithm for bitcoin price forecasting,” in the 13th multidisciplinary academic conference in Prague, 2018, 171–180 (2018).