**Forecasting Drought Vulnerability in Northwestern Bangladesh Using Machine Learning and Deep Learning Approaches**

Prosenjit Basak Arka1, Md. Fakrul Islam1, Mahfuzer Rohman1, Md Sabbir Hossain1, Md Jamal Uddin1,2\*

Prosenjit Basak Arka

[prosenjit03@student.sust.edu](mailto:prosenjit03@student.sust.edu)

Md. Fakrul Islam

[fakrul50@student.sust.edu](mailto:fakrul50@student.sust.edu)

Mahfuzer Rohman

[mahfuzer23@student.sust.edu](mailto:mahfuzer23@student.sust.edu)

Md Sabbir Hossain

[bmsabbirhossainsakib@gmail.com](mailto:bmsabbirhossainsakib@gmail.com)

Md Jamal Uddin

[jamal-sta@sust.edu](mailto:jamal-sta@sust.edu)

1. Biostatistics, Epidemiology and Public Health Research Team, Department of Statistics, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh

2. Faculty of Graduate Education, Daffodil International University, Dhaka, Bangladesh

\* Correspondence to: Md Jamal Uddin, PhD, Professor, Department of Statistics, Shahjalal University of Science and Technology, Sylhet, Bangladesh. Email: [jamal-sta@sust.edu](mailto:jamal-sta@sust.edu)

**Abstract**

**Background**

Agriculture, a cornerstone of Bangladesh's economy, is highly susceptible to weather extremes and water shortages, notably drought. The northwestern region, with its semi-arid climate, faces significant drought vulnerability, impacting millions. The objective of this study is to evaluate and compare advanced machine learning and deep learning-based models for forecasting drought vulnerability in this region, addressing a critical gap in localized drought prediction.

**Methods**

This study focuses on the northwestern region of Bangladesh, specifically using meteorological data from the Rangpur division. The primary data source is the daily rainfall time series from 1981 to 2021, provided by the Bangladesh Meteorological Department. The Standardized Precipitation Index (SPI) is employed to assess drought, with calculations facilitated by the Standardized Drought Analysis Toolbox (SDAT) for enhanced accuracy. Machine learning models, including K-Nearest Neighbors, Random Forest, XGBoost, AdaBoost, CatBoost, and Gradient Boosting, along with deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, are utilized to forecast drought vulnerability. The study emphasizes ensemble methods and uncertainty estimation techniques like Bayesian deep learning and Monte Carlo dropout to improve prediction accuracy and reliability. Data quality is ensured through rigorous testing, including the Augmented Dickey-Fuller test for stationarity. The performance of the model is assessed using metrics such as R², RMSE, MAE, and Pearson's correlation coefficient, alongside visual analyses to assess prediction accuracy and biases

**Results**

The study discovered no specific trend in any of the SPI time series. Among the models, CNN and LSTM had the best fit for SPI-3, SPI6 and SPI-12. Particularly, CNN had R2 value of 58%, 57% and 56% for SPI-3, SPI-6 and SPI-12, respectively. Like R2, CNN also had superior performance metrics. In SPI-3, CNN had values of 69%, 58% and 85% for MAE, NSE and index of agreement, respectively. Similar results were found for SPI-6 and SPI-12 as well. Boosting algorithms such as XGBoost, Gradient Boosting also generally scored high on several performance metrics. On the other hand. KNN and Adaptive Boosting had consistently the worst fit among all the models. KNN had the lowest R2 values of 39% and 37% in SPI-3 and SPI-6, respectively. Adaptive Boosting had the worst fit in SPI-3 and had the lowest R2 value of 30%.

**Conclusion**

In the study, deep learning models have achieved better results than conventional machine learning models. Specifically, CNN and LSTM outperformed all other models in almost all performance metrics for all of SPI-3, SPI-6 and SPI-12. The study suggests using deep learning-based models like CNN and LSTM for predicting drought vulnerability.

**Keywords:** Drought Forecasting, Machine Learning, Deep Learning, Bangladesh, CNN, LSTM

1. **Introduction**

Agriculture plays a crucial role in both national and subnational economies, yet it is highly vulnerable to extreme weather events and water scarcity.[1]. The agricultural drought is linked with the characteristics of the soil. It has become one of the most serious worldwide issues because of insufficient water supply [2]. According to recent research, Both human activity and climate change are projected to increase the frequency and intensity of droughts [3-5]. In the last few years, Bangladesh has experienced a sharp increase in the frequency of droughts [6]. Bangladesh relies heavily on agriculture for its economy and livelihoods. This sector frequently suffers greatly because of natural disasters such as drought and flooding, which are causing significant crop loss [7]. Several studies exist examining the effect of droughts on agriculture, the economy, food production, and society [8-9]. The northwestern area of Bangladesh, which has a semi-arid environment, is particularly vulnerable to drought, affecting the livelihoods of millions and mandating adequate water management and sustainable farming techniques [10,11].

Located in tropical South Asia, Bangladesh contends with recurrent droughts marked by insufficient precipitation [21]. During the period spanning 1960 to 1991, Bangladesh experienced nineteen droughts [25]. Prolonged rainfall deficits induce enduring water shortages, adversely impacting agriculture, environment, and public health [22]. Agricultural losses from drought outweigh those from other natural disasters [23]. Some regions depend entirely on rainfall for agricultural stages, exacerbating the impact. Approximately 2.5 million hectares are impacted by drought annually during the wet season, while around 1.2 million hectares are affected each year in the dry season [24]. Due to its significant agricultural sector and high vulnerability to drought, developing accurate drought forecasting capabilities is crucial for Bangladesh to enhance preparedness, mitigate impacts, and ensure food and water security.

Accurate and early forecasting of drought risk is critical to the goal of proactive disaster preparedness, resource allocation, and long-term planning [12]. Traditional statistical approaches have always served this role, but their limitations in capturing the complicated non-linear connections inherent in drought data have forced academics to investigate novel options [13]. Deep learning and machine learning techniques have transformed environmental and climate forecasting activities in recent years [14, 15], marking the beginning of a new age in advanced drought vulnerability assessment and forecasting**.**

Advanced machine learning techniques like Random Forests (RF), decision trees, and Support vector machines (SVM) have become promising alternatives for drought forecasting, surpassing traditional statistical approaches [16]. On the other hand, the questor's more nuanced, geographically, and temporally aware predictions have spurred the adoption of deep learning models such as Convolutional Neural Networks (CNNs), Recurrent and Neural Networks (RNNs) [17, 18]. These sophisticated models have shown extraordinary skill in capturing complicated spatiotemporal patterns inherent in climate data, making them perfect candidates for precision drought vulnerability predictions.

Ensemble approaches have emerged as attractive tools for further improving prediction accuracy as the academic community goes deeper into the field of drought forecasting [19]. Ensemble approaches combine the benefits of numerous models, reducing the flaws of individual models and improving their ability to predict outcomes. Furthermore, the increased emphasis on uncertainty estimation techniques such as Bayesian deep learning and Monte Carlo dropout have permitted the measurement of prediction uncertainty, instilling trust in decision-making processes [20].

While the current body of literature gives information on the prospect of deep learning methods and machine learning models for assessing drought susceptibility, research relevant to Bangladesh's northwestern area remains scarce. Recognizing the distinct meteorological circumstances and socioeconomic elements that determine drought risk in this region, it is critical to design a specialized forecasting methodology that takes regional differences into consideration. In light of this gap, this study aims to conduct a thorough assessment and Comparison of cutting-edge deep learning techniques and machine learning techniques for predicting drought vulnerability in northwestern Bangladesh. The insights garnered from this research can inform drought preparedness strategies, guiding resource allocation decisions to enhance the well-being and resilience of the northwestern region's populace in Bangladesh.

1. **Methodology**

**2.1 Study Area**

Bangladesh is a country in Southeast Asia that lies between latitudes 20°34' - 26°38'N and longitudes 88°01' - 92°41'E. The country's hydro-geologic environment is diverse and complicated, with few mountain zones and flat, low-lying plains [26]. There is total 35 meteorological stations available in Bangladesh. In this study, to analyze the droughts, northwestern region specifically rangpur division meteorological stations were used.

**2.2 Data Collection and Preparation**

A popular tool for evaluating drought is the Standardized Precipitation Index (SPI). The World Meteorological Organization (WMO) has recognized it as the instrument that national meteorological and hydrological services should use to monitor drought. Since SPI computation requires only data on rainfall, the monthly dataset of rainfall time series has been examined. The Bangladesh Meteorological Department provided the daily rainfall time series data for 1981–2021. SPI is a frequently used method for evaluating meteorological drought severity, but it has limitations because it assumes that an appropriate probability distribution function exists for precipitation data modeling. Furthermore, it is impossible to compare the values of various indicators directly because they have distinct scales and ranges. A solution to these problems is offered in the form of the Standardized Drought Analysis Toolbox (SDAT). Nonparametric univariate and multivariate approaches can be used to assess different climate and land-surface parameters because it provides a broad foundation for developing standardized indices without the need for representative parametric distributions [27]. Consequently, we use SDAT to calculate SPI (SDAT-SPI) in this study.

**SPI Scales Used in this Study**

• 3-month SPI.

• 6-month SPI.

• 12-month SPI.

The data was cleaned and quality validated prior to usage in the modeling procedure. The Nearest Neighbor Method has been used to impute several missing values that were present in the observed data. After that, the data has been converted into monthly data on rainfall.

**2.3 Standardized Precipitation Index (SPI) & Standardized Drought Analysis Toolbox (SDAT)**

While calculating the SPI has traditionally involved fitting a gamma probability distribution function to precipitation data, this method may not be appropriate. The necessity for a comprehensive approach that considers multiple indicators is further highlighted by the shortcomings of the present drought indicators, which include statistical incomparability and temporal inconsistency. A comprehensive methodological framework for creating nonparametric standardized indices across a variety of ground-surface and meteorological data is provided by the Standardized Drought Analysis Toolbox (SDAT). The creation of multivariate drought indices based on many variables is made possible by this nonparametric approach, which can also increase temporal and spatial consistency. The empirical Gringorten plotting position method has been proposed for estimating marginal probability distributions of precipitation and other climatic variables. This approach uses a particular technique to determine the data's rank-based probability distribution, which can be useful in estimating probability of specific precipitation events occurring.

|  |  |  |
| --- | --- | --- |
|  |  | (1) |

Here, in Eq. 1, the sample size is represented by n, and i is the sequential ranking, from minimum to maximum, of non-zero precipitation data. The corresponding empirical probability is denoted by  . This empirical method eliminates the need to derive parametric probabilities. It is possible to transform the results obtained from the equation into a Standardized Index (SI).

|  |  |  |
| --- | --- | --- |
|  |  | (2) |

where represents the standard normal distribution function, and represents the probability derived from equation. Percentiles can also be standardized by utilizing a commonly accepted approximation formula of the previous equation. This can be useful in situations where percentile values need to be compared across different datasets or variables with varying scales and ranges.

|  |  |  |
| --- | --- | --- |
|  |  | (3) |

where and

|  |  |  |
| --- | --- | --- |
|  |  | (4) |

**2.4 Interpretation of SPI values**

Wet weather is indicated by positive values, whilst dry weather is shown by negative values. Very dry weather is defined as anything below -2, whereas extremely wet weather is defined as anything above 2. The overall weather range of a region is between -2 and +2.

**2.5 Data quality check**

An additive decomposition approach was employed to break down the data into trends, seasonal patterns, and irregular components. Understanding the essential structure of the data is essential. To make the data stationary, differencing was applied. To ensure the dependability of the analysis, the stationary nature of the time series data was evaluated using the Augmented Dickey-Fuller test.

Next, deep learning and machine learning techniques were fitted to data. Several measures were computed to assess the performance of these techniques. This comprehensive process allowed analyzing and predicting patterns effectively, providing a robust framework for the study.

**2.6 Theoretical Overview of ML and DL models**

This subsection offers a brief overview of several deep learning (DL) and machine learning (ML) techniques.

**2.6.1 Random Forest**

The Random Forest (RF) model employs an ensemble learning technique combined with a bagging approach, utilizing decision trees to address regression problems. In RF, each node in the data is split randomly based on a selection of the most influential input variables [28]. This improves the learning process, improving prediction accuracy and preventing overfitting. Here are the steps to build an RF model:

1. Select K data points are selected at random from the training dataset.
2. Build the decision tree using the data from step 1.
3. Decide on the number of decision trees (ntrees) to build.
4. Repeat steps 1 and 2 for each tree.
5. Combine the predictions from all the trees to forecast the multi-scaler SPI.

**2.6.2 Convolutional Neural Network**

Convolutional Neural Networks (CNNs) is a deep learning model that utilize convolution operations to process inputs during both training and error propagation. CNNs consist of trainable layers, with a key advantage being their ability to learn directly from input features. They are well-suited for scenarios where dependencies exist within the input data, helping to reduce computational costs and mitigate overfitting. A Convolutional Neural Network (CNN) is structured with pooling layers, convolutional layers, and fully connected layers. The convolutional layers are the heart of a CNN; they use filters to generate feature maps from the input data. Then, pooling layers condense these feature maps, creating smaller, more manageable matrices. Finally, the fully connected layer operates much like a traditional multilayer perceptron neural network. The significant strength of CNNs lies in their capacity to automatically identify important features within the data. This makes them particularly well-suited for data where elements are interconnected and for handling inputs with a large number of dimensions. [29].

**2.6.3 K-Nearest Neighbors**

K-Nearest Neighbors (KNN) is a versatile, non-parametric machine learning method used for applications involving regression and classification. Rooted in data mining principles, it proves particularly powerful for classification. KNN operates by using a labeled training dataset to assign categories to new, unknown data points based on the categories of their closest neighbors in the training data. Evaluating any method involves three key aspects: ease of output interpretation, calculation time, and predictive power. In KNN, an object's class is assigned based on the majority class among its closest neighbors. The parameter *k* is a positive integer indicating how many nearest neighbors are taken into account. Typically, is small, if = 1 the object is classified as its nearest neighbor. The selection of *k* plays a notable effect on achieving the model's performance. If is too small, prediction errors increase because the model may be too sensitive to noise in the data. Conversely, too many neighbors (a high ) value) can lead to overfitting, where the model becomes too complex and less generalizable. Finding the optimal is crucial for achieving the best model performance [30].

**2.6.4 Adaptive Boosting**

Adaptive Boosting (AdaBoost) is an effective machine learning algorithm used for both classification and regression. It strengthens model performance by combining several weak learners into a single, robust classifier. The model undergoes iterative training, adjusting the weights of incorrectly classified samples to focus more on them in subsequent iterations [31].

To evaluate the AdaBoost model, three major aspects are considered: ease of output interpretation, calculation time, and predictive power.

1. **Ease of Output Interpretation**: AdaBoost provides a clear and interpretable output by combining the predictions of multiple weak classifiers. Each weak classifier contributes to the final decision based on its accuracy, making the overall decision easy to understand.
2. **Calculation Time**: AdaBoost may require significant computational resources, particularly when applied to large datasets or run with many iterations. However, its efficiency in improving classification accuracy often outweighs the additional computation time.
3. **Predictive Power**: AdaBoost’s key advantage is its ability to improve the predictive performance of weak learners. By focusing on misclassified samples, it significantly enhances the model's accuracy. However, it is essential to avoid overfitting by carefully tuning the number of iterations and other parameters.

**2.6.5 XGBoost**

XGBoost is a powerful gradient boosting algorithm designed for classification and regression tasks, recognized for its speed, accuracy, and scalability. XGBoost builds a model in a stage-wise fashion, adding new trees to correct the mistakes of previously built models. This iterative approach helps improve the model's performance. Key aspects to evaluate XGBoost include ease of output interpretation, calculation time, and predictive power. XGBoost's outputs can be interpreted by examining feature importance scores, which illustrate the impact of each feature on the model's predictions. Calculation time can be optimized with parallel processing and tree pruning techniques, making XGBoost faster than many other algorithms. Predictive power is one of XGBoost's strengths, as it tends to provide high accuracy and robust predictions. It includes regularization parameters to prevent overfitting, improving the model's generalization to new data. Hyperparameters such as learning rate, max depth, and the number of estimators must be carefully tuned to achieve the best performance [32].

**2.6.6 CatBoost**

CatBoost is a gradient boosting-based machine learning algorithm and is particularly effective for handling categorical data. It is applied to both classification and regression problems. CatBoost uses a training set to build an ensemble of decision trees, which are combined to make predictions. This algorithm is notable for its capability to manage categorical features without the need for extensive preprocessing, as well as its resistance to overfitting. To evaluate a CatBoost model, there are three major aspects: ease of output interpretation, calculation time, and predictive power. CatBoost provides easy-to-interpret outputs by showing feature importance and contributions to the prediction. It is optimized for fast training and prediction times, even with large datasets and numerous categorical features. The CatBoost model's predictive power is enhanced through several techniques, such as ordered boosting, which helps prevent overfitting, and symmetric trees, which improve training speed and accuracy. Evaluating the model involves measuring its performance using metrics like mean squared error or accuracy for classification tasks for regression tasks. Cross-validation is also employed to verify the model’s ability to generalize to new data, offering a dependable measure of its predictive power [33].

**2.6.7 Gradient Boosting**

Gradient Boosting is an effective machine learning technique applied to classification and regression tasks. It constructs models in a sequence, where each model aims to correct the errors made by its predecessor. This method leverages the strengths of several weak learners, usually decision trees, to build a robust predictive model. When evaluating the Gradient Boosting model, three key factors are considered: ease of output interpretation, calculation time, and predictive power. Known for its excellent accuracy, Gradient Boosting is a favored choice for handling complex datasets. However, the interpretability of its outputs can be challenging because the final model is an ensemble of many decision trees. The calculation time for Gradient Boosting can be significant, especially for large datasets, because it builds multiple trees sequentially. Each iteration involves computing residuals and fitting a new tree, which can be computationally intensive. In Gradient Boosting, hyperparameters such as the number of trees, learning rate, and tree depth play a vital role in model performance. A small learning rate with many trees typically improves accuracy but increases computation time. Conversely, a higher learning rate with fewer trees may lead to faster training but can increase the risk of overfitting. Proper tuning of these hyperparameters is essential to balance efficiency and accuracy [34].

**2.6.8 Long Short-Term Memory**

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture developed to address the issues of vanishing and exploding gradients in traditional RNNs. LSTM is well-suited for learning from experience, classifying processes, and predicting time series when there are unspecified time delays and boundaries between significant events. It utilizes a memory component known as the cell state to capture long-term dependencies. LSTM has a chain-like structure with four gates, each with a different function: forget, input, candidate, and output. Forget, input, and output gates use sigmoid functions, while candidate and output gates use tanh functions. Logistic functions are calculated using weights and biases to activate neurons and normalize inputs. Each neuron in the hidden layer of a recurrent neural network is equipped with an LSTM unit and goes through multiple states during the prediction process. LSTM is particularly effective in learning from experience and classifying processes when there are unknown time delays and bounds between significant events [35]. These models have demonstrated their effectiveness in predicting environmental management, droughts, rainfall, solar index values, and soil moisture levels. For more detailed studies on these models, readers can refer to the articles 36 – 39.

**2.7 Performance Evaluation**

The performance of each model will be evaluated using various metrics suitable for the regression task. Common metrics such as Coefficient of Determination (R^2), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Index of Agreement (d), Pearson's Correlation Coefficient (r) and Nash-Sutcliffe Efficiency (NSE) are calculated to assess model accuracy and predictive ability. Additionally, visualizations such as scatter plots and residual plots are generated to provide insight into model predictions and potential biases.

r =

MAE =

RMSE =

d = 1 -

NSE = 1 -

RAE =

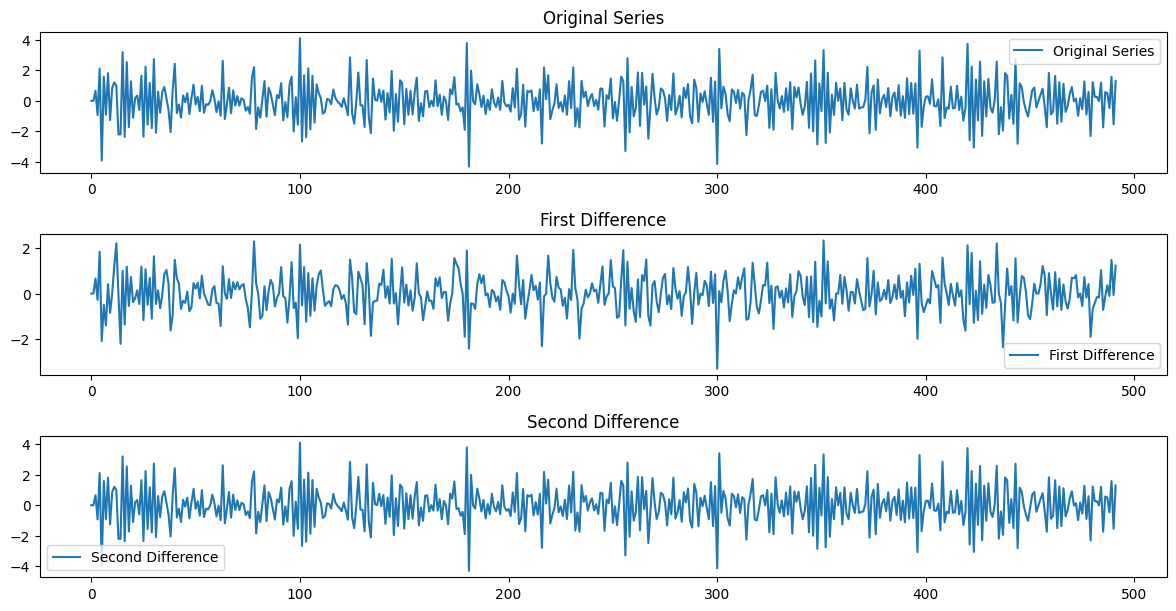
Here, SPIA is an actual or observed value, SPIP is predicted value, and and are the mean values of observed and predicted samples, and N represents the total number of data points.

1. **Results**

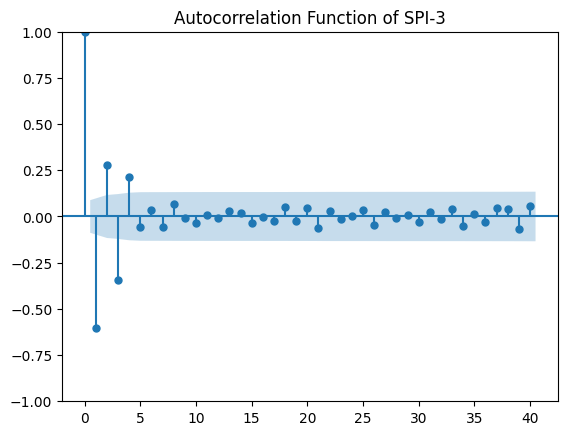
**3.1 Trend of Standardized Precipitation Index (SPI-3) of Rangpur Division From 1981 to**

**2021**

**Figure 3.1(a)** represents the trend of original, first and second difference SPI-3 from 1981 to 2021. The study did not find any specific trend in any of the time indicators.



**Figure 3.1(a)**: Trend of Original, First and Second Difference of SPI-3 of Rangpur Division from 1981 to 2021.



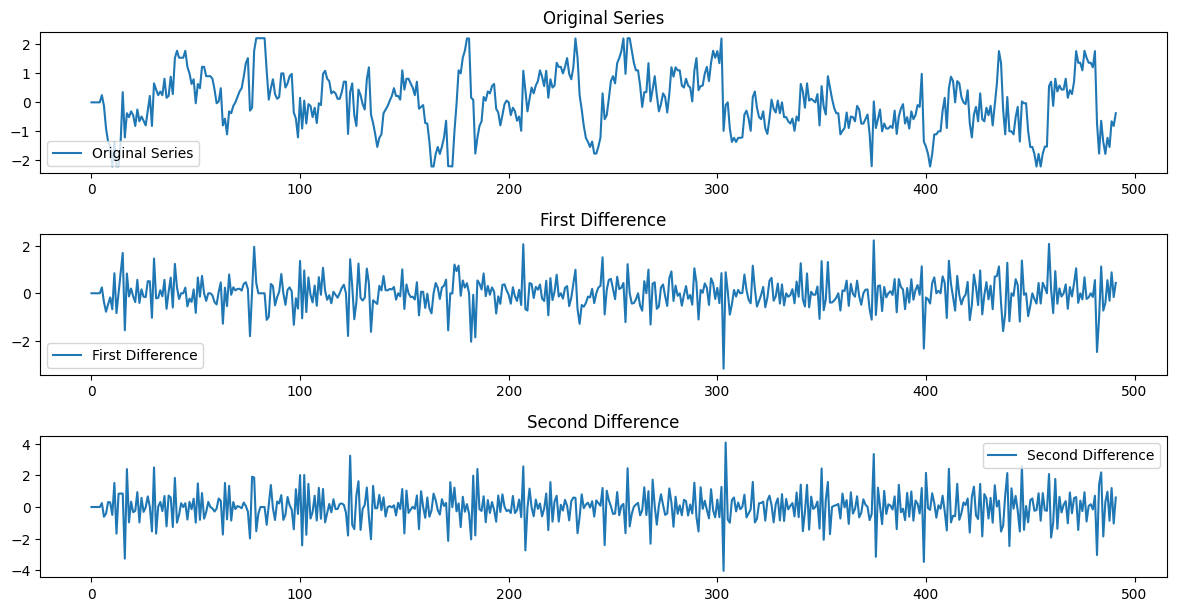
**Figure 3.1(b):** PACF of SPI-3 of Rangpur Division from 1981 to 2021.

**Figure 3.1(b)** represents partial auto-correlation function (PACF) of second difference SPI-3. In Durbin-Watson test, the value of the test statistic was 3.21, indicating no autocorrelation among the terms.

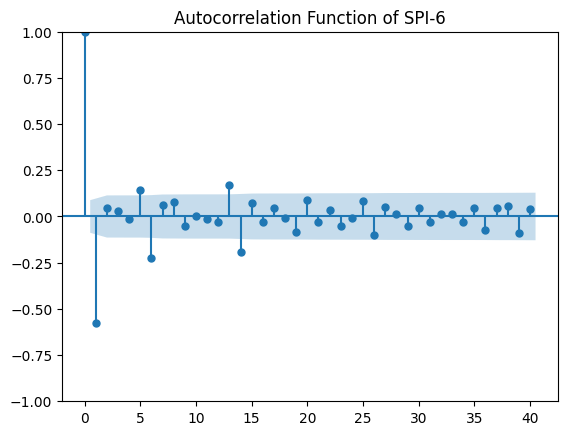
The first difference SPI-3 was also found to be a stationary time series data. The series was significant (Test Statistic = -10.36; p-value < 0.05) in augmented Dickey-Fuller test indicating stationarity.

**3.2 Trend of Standardized Precipitation Index (SPI-6) of Rangpur Division From 1981 to**

**2021Figure 3.2(a)** shows the trend of original SPI-6 and first difference SPI-6 from 1981 to 2021. SPI-6 did not have any specific trend in either the original or the first difference.



**Figure 3.2(a):** Trend of Original, First and Second Difference of SPI-6 of Rangpur Division from 1981 to 2021.



**Figure 3.2(b):** PACF of SPI-6 of Rangpur Division from 1981 to 2021.

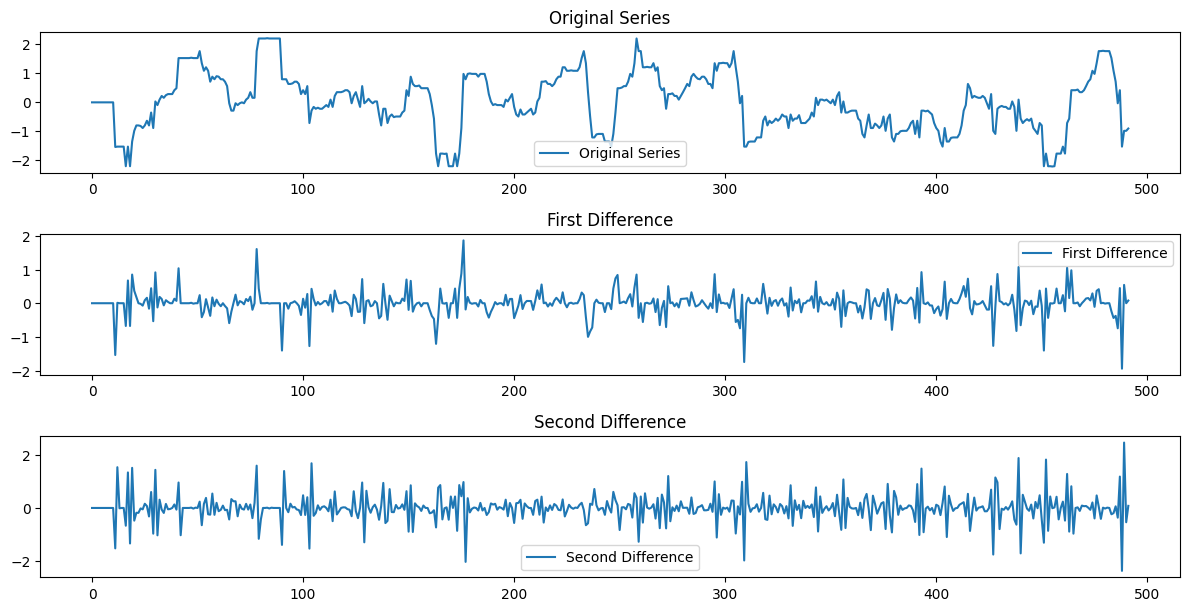
There was no autocorrelation among the terms of SPI-6 as indicated by the PACF plot in **Figure 3.2(b)**. In the Durbin-Watson test, the test statistic was 3.15.

The time series of SPI-6 was checked for stationary using augmented Dickey-Fuller test. The series was significant (Test Statistic = -13.20; p-value < 0.05) in the test indicating stationarity.

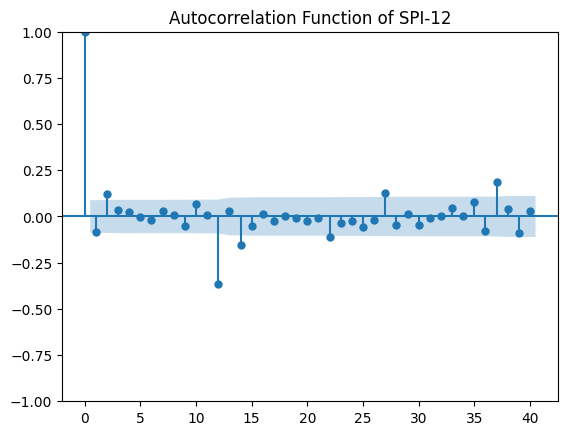
**3.3 Trend of Standardized Precipitation Index (SPI-12) of Rangpur Division From 1981 to**

**2021**

**Figure 3.3(a)** illustrates the trend of original SPI-12 and first difference SPI-12 of Rangpur division from 1981 to 2021. SPI-12 did not have any specific trend in either the original or the first difference.



**Figure 3.3(a):** Trend of Original, First and Second Difference of SPI-12 of Rangpur Division from 1981 to 2021.

****

**Figure 3.3(b):** PACF of SPI-12 of Rangpur Division from 1981 to 2021.

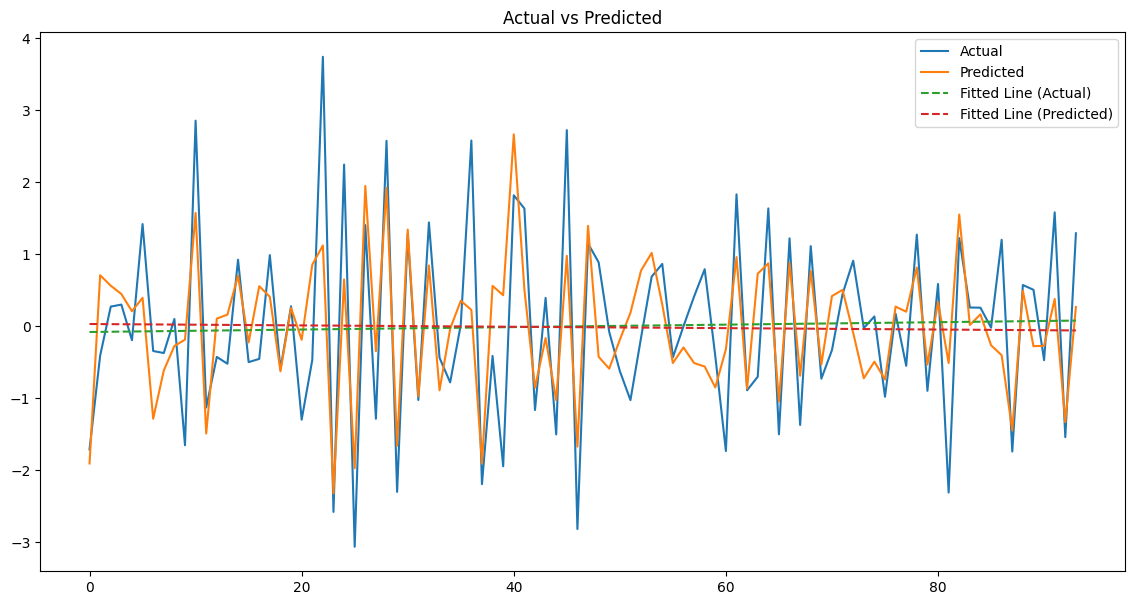
**Figure 3.3(b)** shows the PACF of SPI-12. According to Durbin-Watson test, the value of the test statistic was 3.19, indicating no autocorrelation between the terms of the series.

The first difference SPI-12 was found to be a stationary time series data. The series was significant (Test Statistic = -9.48; p-value < 0.05) in augmented Dickey-Fuller test indicating stationarity.

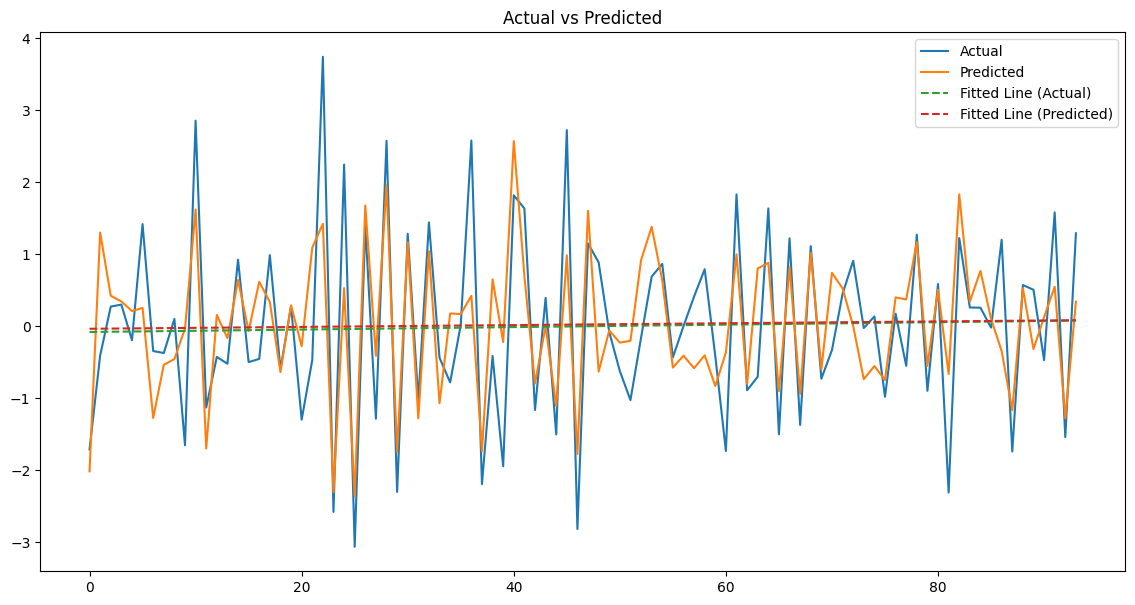
**3.4 Performance Evaluation of Machine Learning and Deep Learning Model for SPI-3**

For SPI-3, LSTM, Random Forest, XGBoost, CatBoost, Gradient Boosting and CNN all fitted well (above 50% in R2 value). Among them CNN had the best fit with R2 value of 58%. On the other hand, KNN and Adaptive Boosting had the worst fit with R2 value of 39% and 43%, respectively.

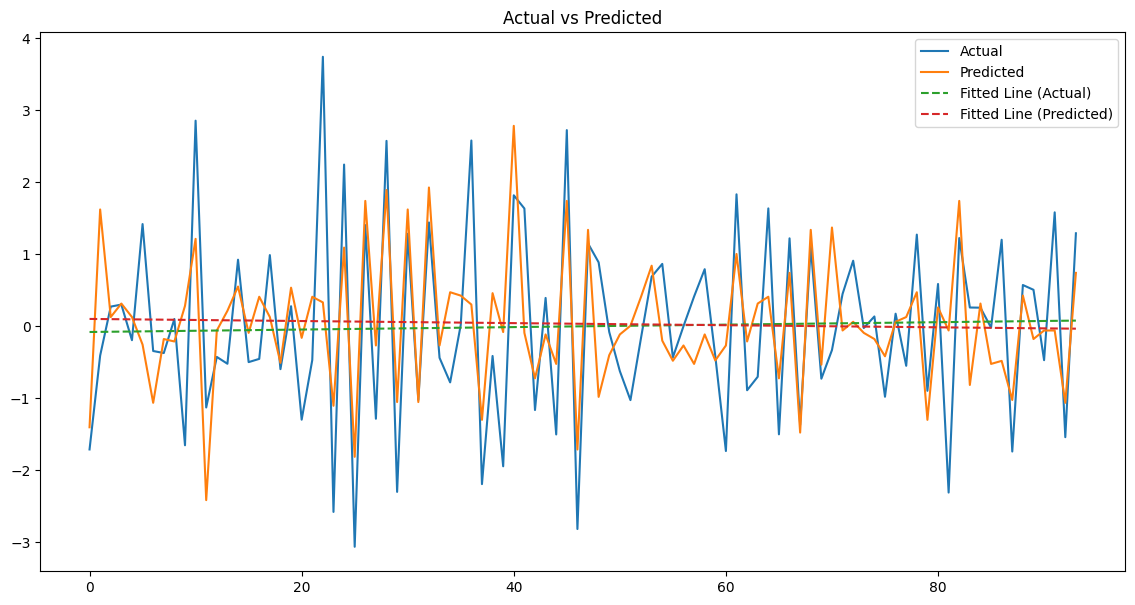
**Figure 3.4** shows the observed vs predicted plots of all the above models



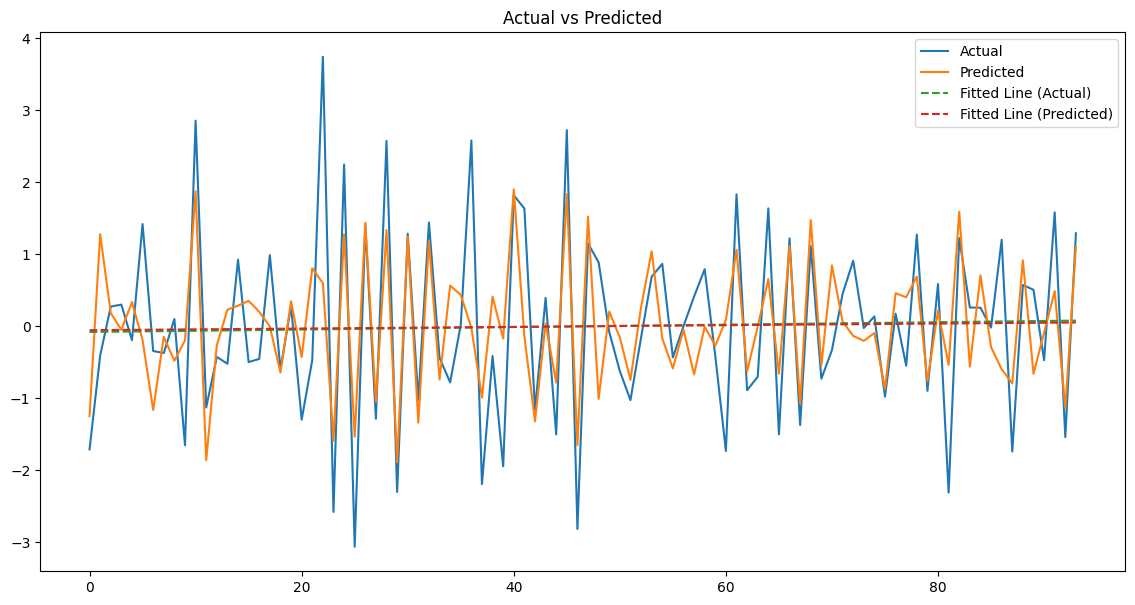
**Figure 3.4(a):** Actual Vs Predicted Plot of LSTM for SPI-3



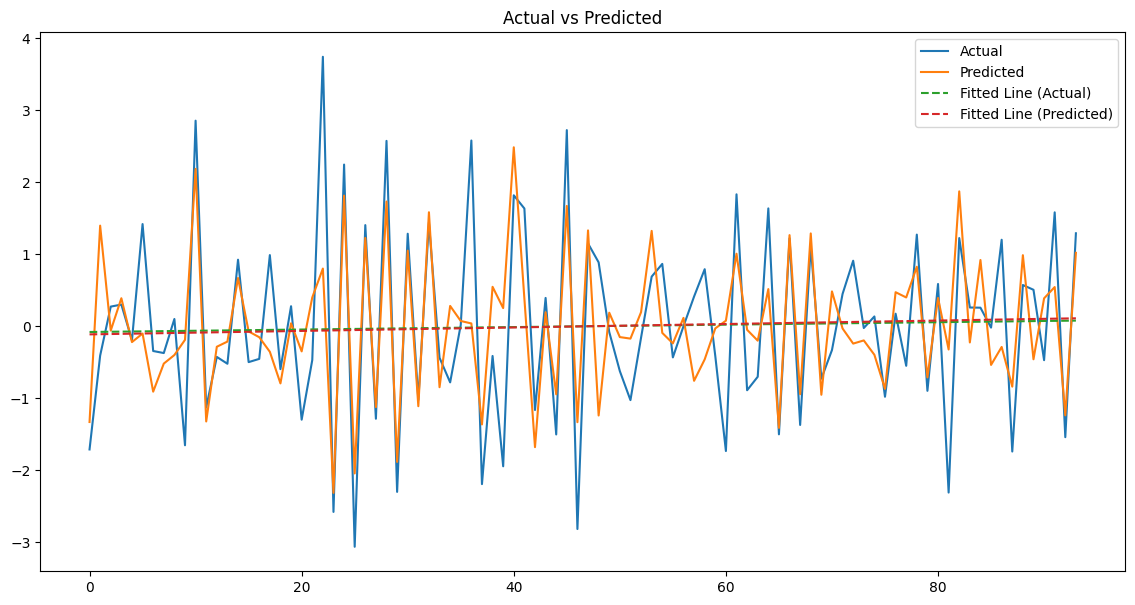
**Figure 3.4(b):** Actual Vs Predicted Plot of CNN for SPI-3



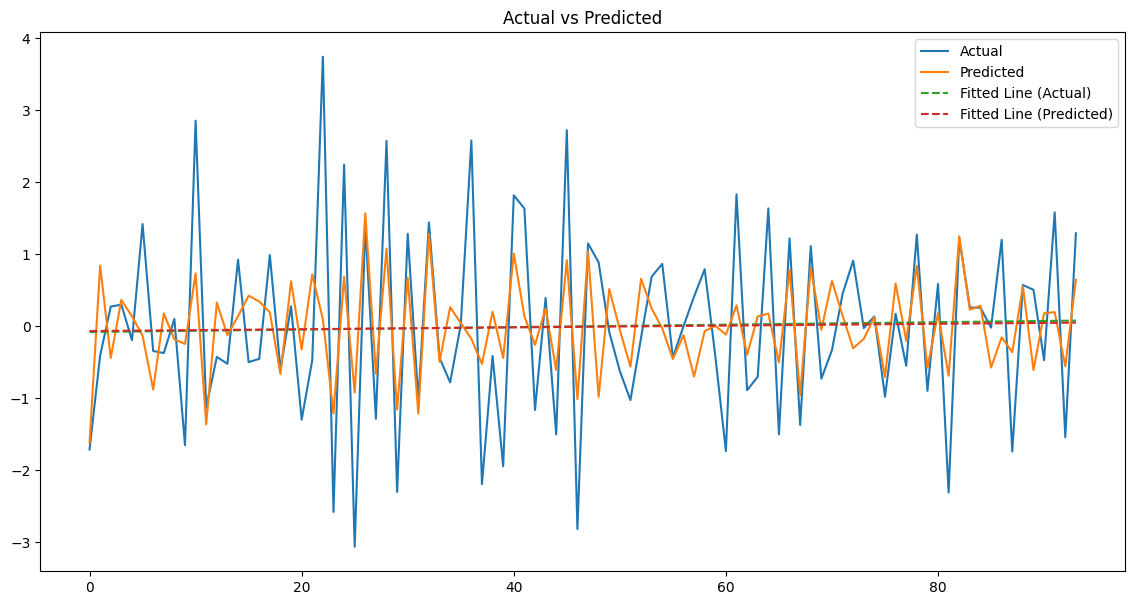
**Figure 3.4(c):** Actual Vs Predicted Plot of Adaptive Boosting for SPI-3



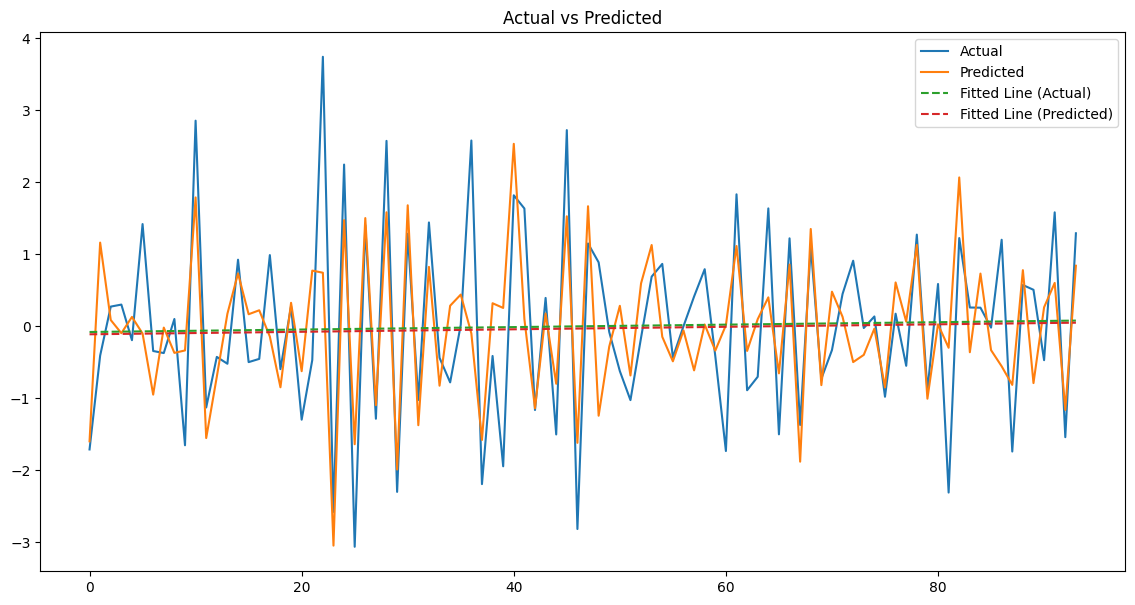
**Figure 3.4(d):** Actual Vs Predicted Plot of CatBoost for SPI-3



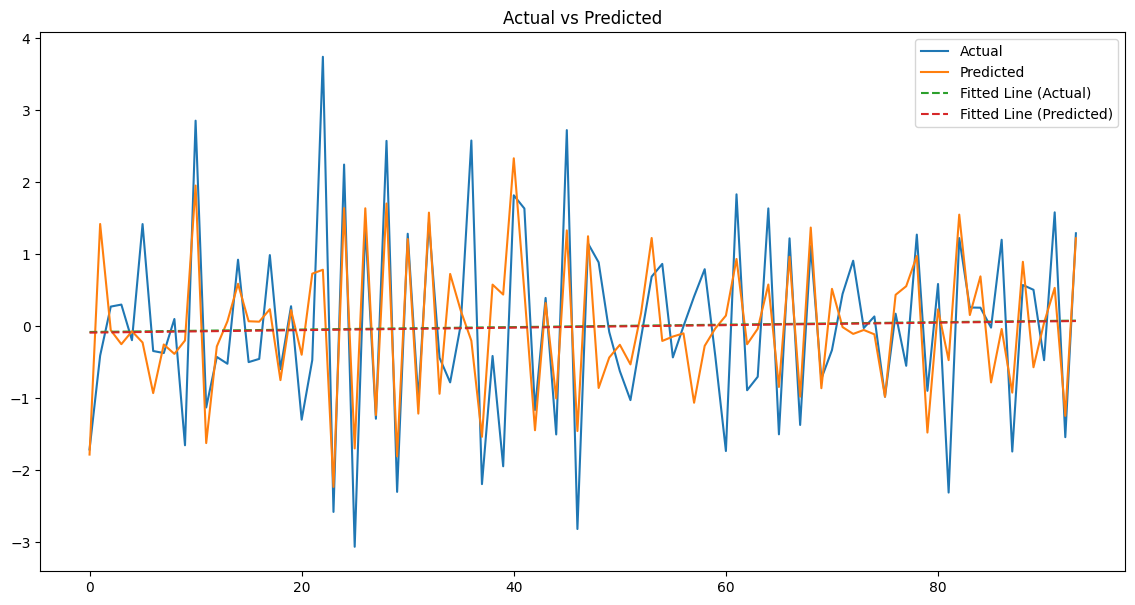
**Figure 3.4(e):** Actual Vs Predicted Plot of Gradient Boosting for SPI-3



**Figure 3.4(f):** Actual Vs Predicted Plot of KNN for SPI-3



**Figure 3.4(g):** Actual Vs Predicted Plot of Random Forest for SPI-3



**Figure 3.4(h):** Actual Vs Predicted Plot of XGBoost for SPI-3

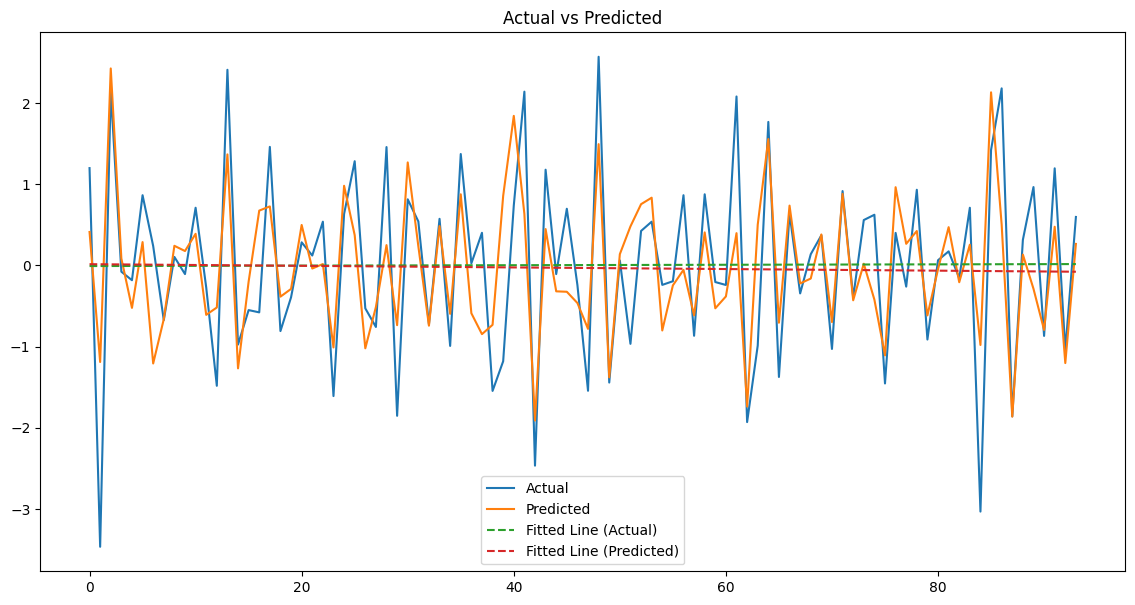
Assessing the other evaluation metrics, Gradient Boosting performed the best in MAE with the lowest value of 68%. LSTM and CNN also performed well with a value of 69%. KNN performed the worst with the highest value of 81%. Similarly, in RMSE, Gradient Boosting achieved the best performance with the lowest value of 68% and KNN performed the worst with the highest value of 105%. In NSE, CNN performed the best with the value of 58%. LSTM and Gradient Boosting also performed quite well with the value of 56% and 55%, respectively. Conversely, KNN performed the worst with a value of 39%. CNN also performed the best in the index of agreement (d) with the value of 85%. Similarly, LSTM, XGBoost and Gradient Boosting performed better than other models with all three having the value of 83%. Conversely, KNN had the worst performance with the value of 69%. In RAE, CNN and Gradient Boosting outperformed all other models with 64%. LSTM and XGBoost also performed quite well with 65% each. KNN, as always, performed the worst with the value of 76%. CNN outperformed all other models in Pearson’s Correlation Coefficient (r) with the value of 76%. LSTM and Gradient Boosting were close seconds with the values of 75% and 74%, respectively. Conversely, KNN had the worst value of 64% [**Table 3.1**].

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Evaluation Metric | LSTM | Random Forest | KNN | Adaptive Boosting | XG Boosting | Cat Boosting | Gradient Boosting | CNN |
| MAE | 0.69 | 0.72 | 0.81 | 0.79 | 0.69 | 0.72 | 0.68 | 0.69 |
| RMSE | 0.88 | 0.93 | 1.05 | 1.00 | 0.91 | 0.93 | 0.68 | 0.86 |
| R2 | 0.56 | 0.51 | 0.39 | 0.43 | 0.53 | 0.51 | 0.55 | 0.58 |
| NSE | 0.56 | 0.51 | 0.39 | 0.43 | 0.53 | 0.51 | 0.55 | 0.58 |
| d | 0.83 | 0.82 | 0.69 | 0.77 | 0.83 | 0.81 | 0.83 | 0.85 |
| RAE | 0.65 | 0.68 | 0.76 | 0.74 | 0.65 | 0.68 | 0.64 | 0.64 |
| r | 0.75 | 0.72 | 0.64 | 0.66 | 0.73 | 0.72 | 0.74 | 0.76 |

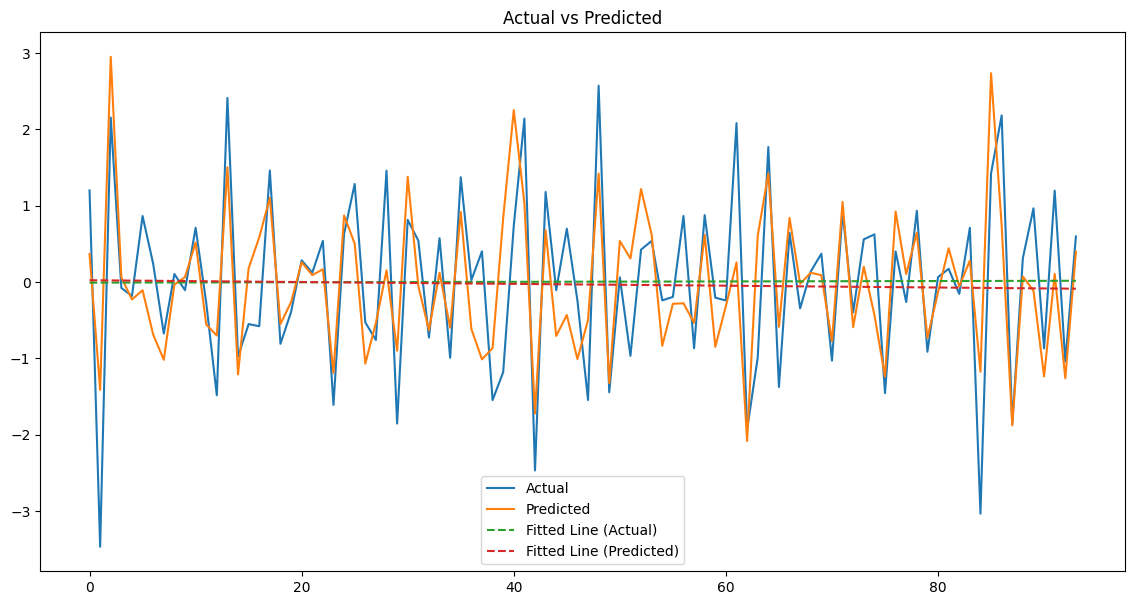
**Table 3.1**: Evaluation Metrics of Machine Learning and Deep Learning Models of SPI-3

**3.5 Performance Evaluation of Machine Learning and Deep Learning Model for SPI-6**

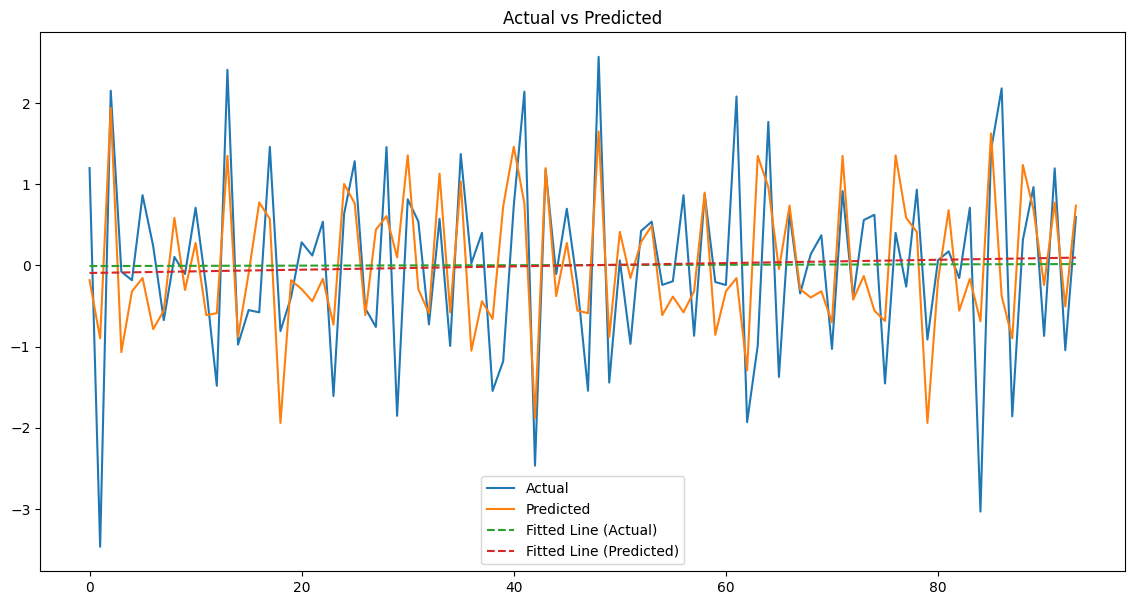
For SPI-6, CNN demonstrated the best fit with R2 value of 57%. LSTM also performed quite well with the R2 value of 56%. On the other hand, KNN and Adaptive Boosting had the worst fit with the R2 value of 37% each. **Figure 3.5** illustrates the actual vs predicted plots of all the models below.



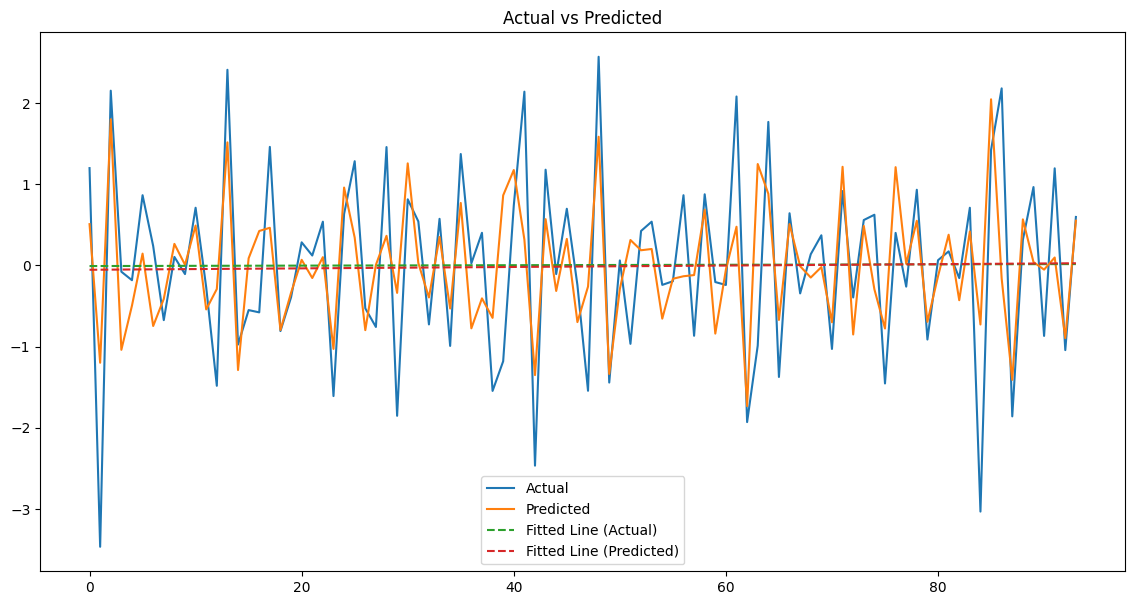
**Figure 3.5(a):** Actual Vs Predicted Plot of CNN for SPI-6



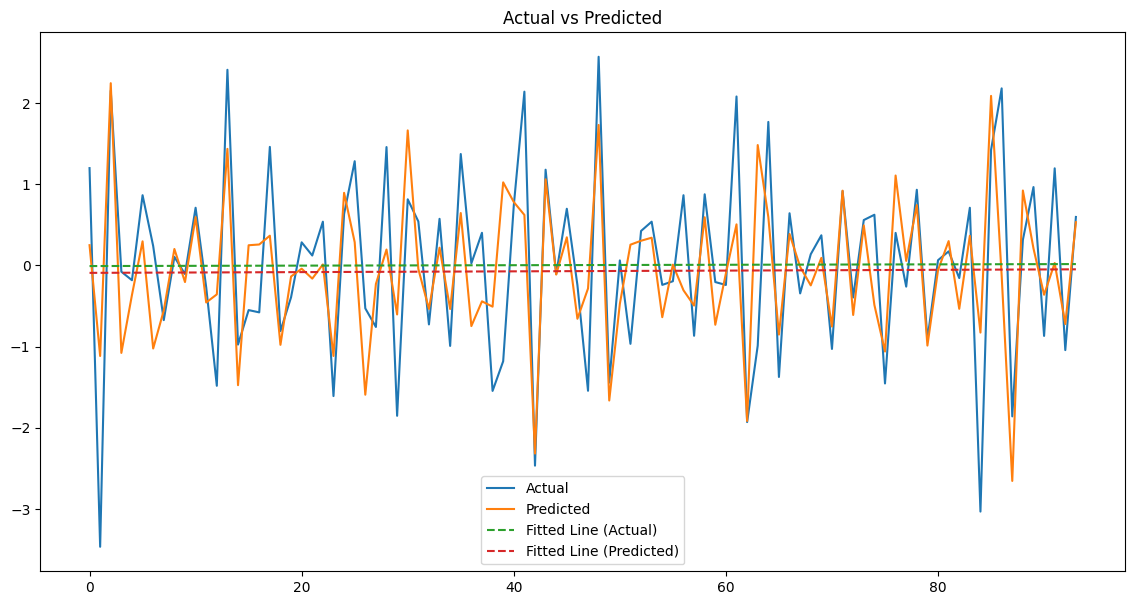
**Figure 3.5(b):** Actual Vs Predicted Plot of LSTM for SPI-6



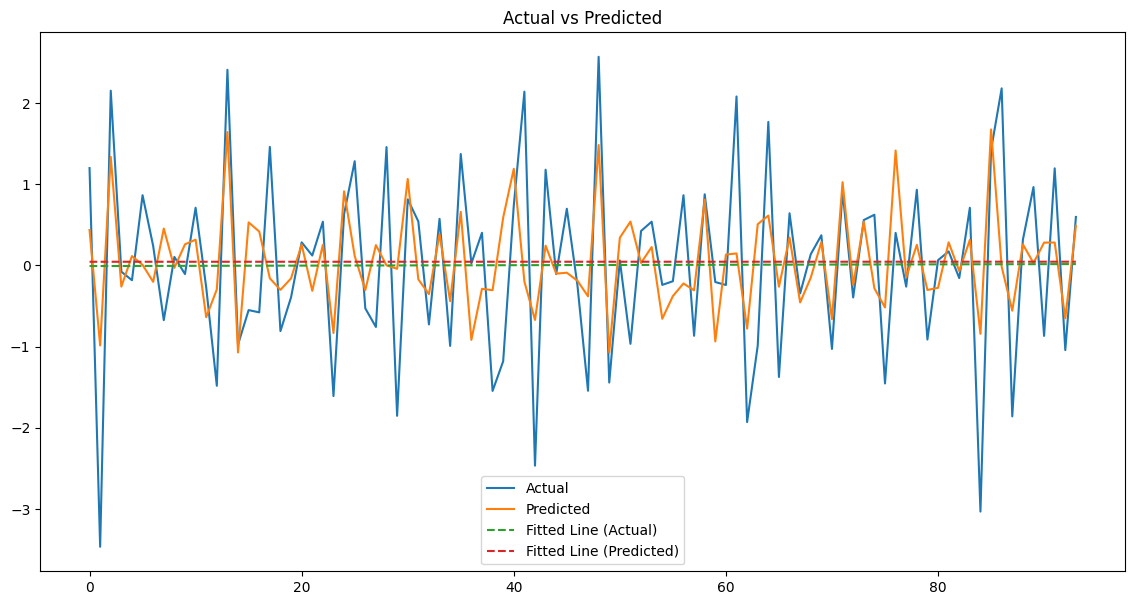
**Figure 3.5(c):** Actual Vs Predicted Plot of Adaptive Boosting for SPI-6



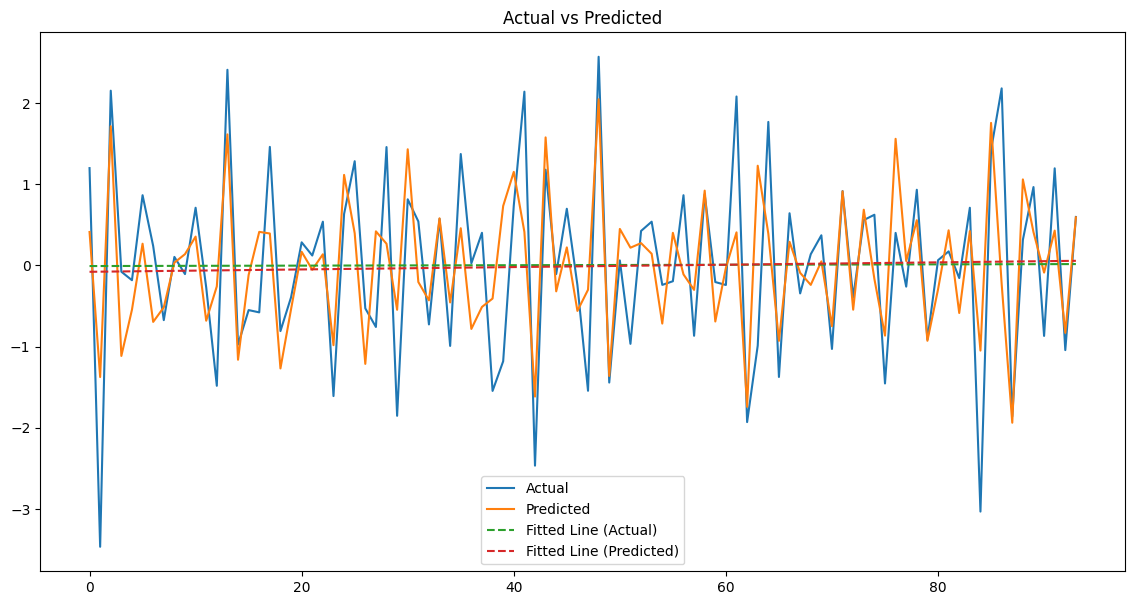
**Figure 3.5(d):** Actual Vs Predicted Plot of CatBoost for SPI-6



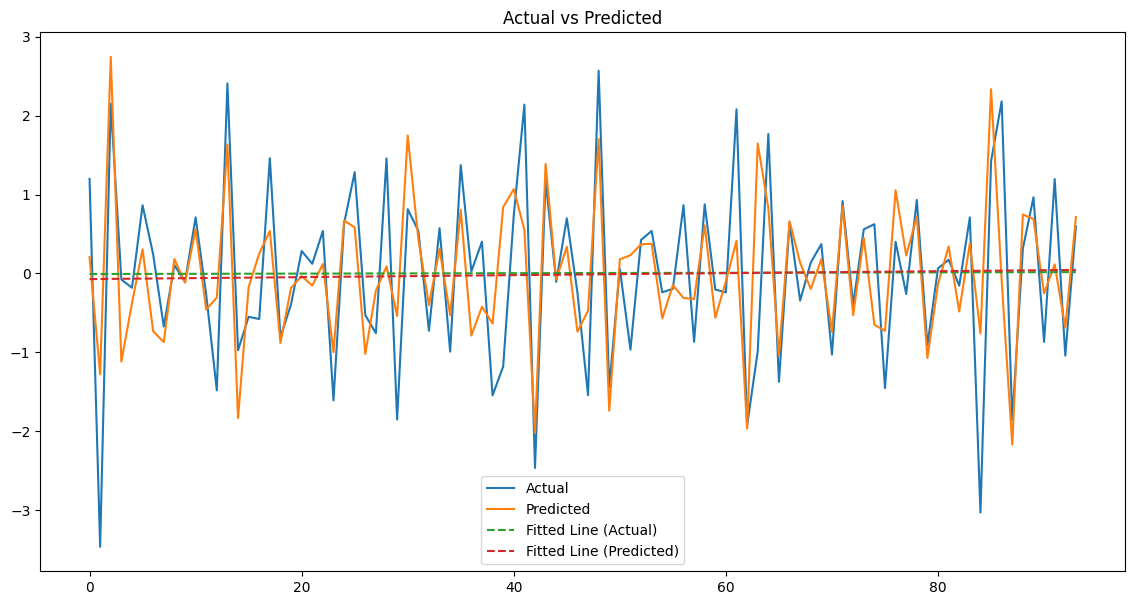
**Figure 3.5(e):** Actual Vs Predicted Plot of Gradient Boosting for SPI-6



**Figure 3.5(f):** Actual Vs Predicted Plot of KNN for SPI-6



**Figure 3.5(g):** Actual Vs Predicted Plot of Random Forest for SPI-6



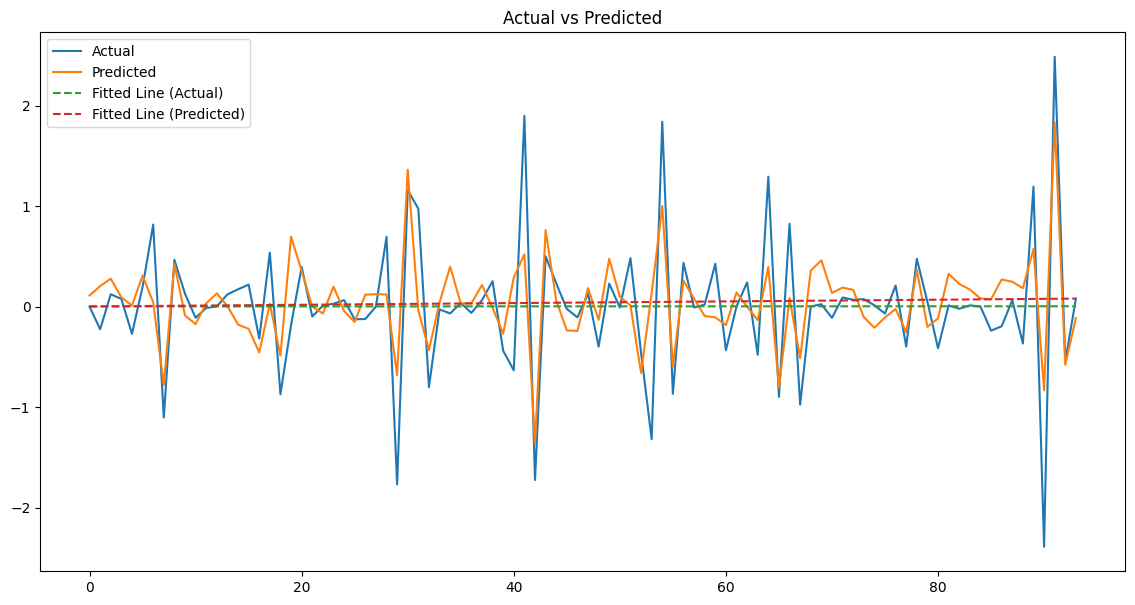
**Figure 3.5(h):** Actual Vs Predicted Plot of XGBoost for SPI-6

CNN performed the best in MAE with lowest value of 57%. LSTM and XGBoost also had comparable performance with a value of 60% each. On the other hand, KNN and Adaptive Boosting had the worst performance with the values of 72% each. In RMSE, CNN also performed the best with a value of 76%, closely followed by LSTM with the value of 77%. KNN and Adaptive Boosting again performed the worst with the values of 92% each. In NSE, CNN again performed the best with a value of 57%, closely followed by LSTM at 56%. Conversely, KNN and Adaptive Boosting performed the worst with the values of 37% each. LSTM outperformed CNN in index of agreement with the value of 85% while CNN had a value of 84%. KNN performed considerably worse than the other models with the value of 69%. In RAE, CNN performed the best with the lowest value of 62%. LSTM and XGBoost also performed quite well with 65% each. KNN and Adaptive Boosting again performed the worst with the values of 78% each. CNN performed the best in Pearson’s Correlation Coefficient with the value of 76%, closely followed by LSTM at 75%. On the other hand, KNN and Adaptive Boosting had the worst performance with the values of 62% each. **Table 3.2** below lists the evaluation metrics of all the models for SPI-6.

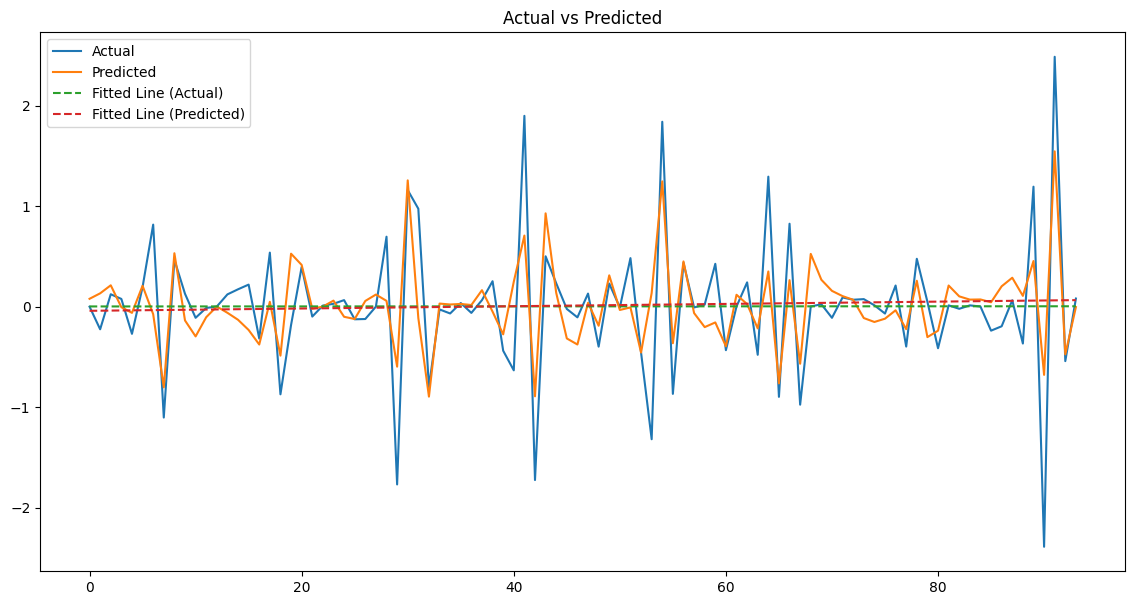
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Evaluation Metric | LSTM | Random Forest | KNN | Adaptive Boosting | XG Boosting | Cat Boosting | Gradient Boosting | CNN |
| MAE | 0.60 | 0.63 | 0.72 | 0.72 | 0.60 | 0.64 | 0.62 | 0.57 |
| RMSE | 0.77 | 0.82 | 0.92 | 0.92 | 0.82 | 0.83 | 0.84 | 0.76 |
| R2 | 0.56 | 0.50 | 0.37 | 0.37 | 0.51 | 0.49 | 0.48 | 0.57 |
| NSE | 0.56 | 0.50 | 0.37 | 0.37 | 0.51 | 0.49 | 0.48 | 0.57 |
| d | 0.85 | 0.81 | 0.69 | 0.76 | 0.82 | 0.79 | 0.81 | 0.84 |
| RAE | 0.65 | 0.69 | 0.78 | 0.78 | 0.65 | 0.70 | 0.68 | 0.62 |
| r | 0.75 | 0.71 | 0.62 | 0.62 | 0.71 | 0.70 | 0.70 | 0.76 |

**Table 3.2:** Evaluation Metrics of Machine Learning and Deep Learning Models of SPI-6

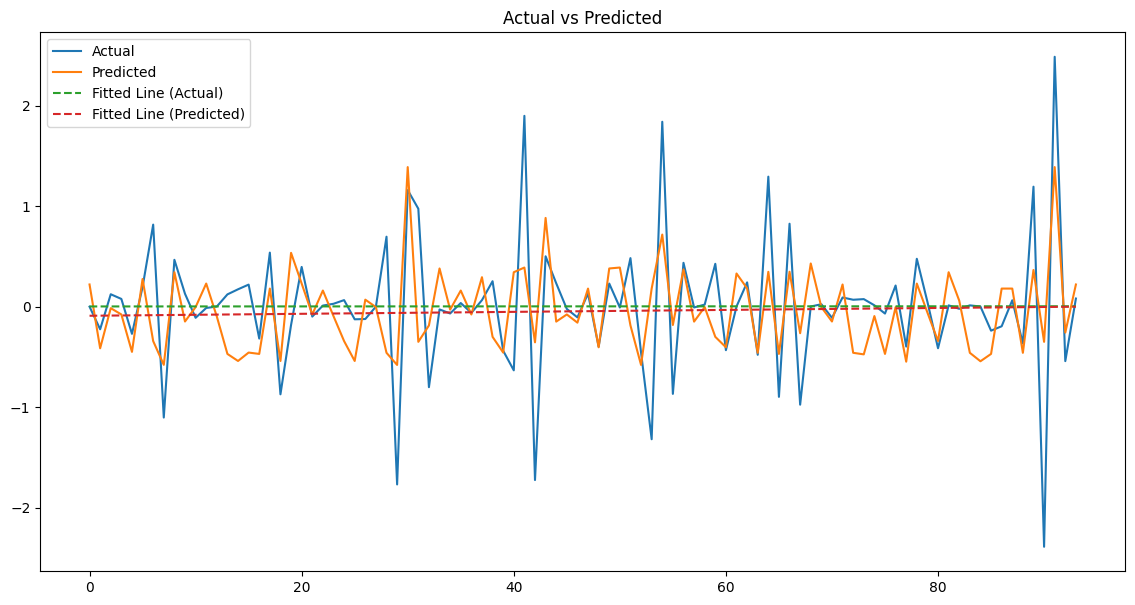
**3.6 Performance Evaluation of Machine Learning and Deep Learning Model for SPI-12**

For SPI-12, CNN and LSTM both had the best fit with R2 values of 56% each. Random Forest and XGBoost also had quite good fit with R2 values of 47% and 46%, respectively. On the other hand, Adaptive Boosting had the worst fit with R2 value of 30%. **Figure 3.6** shows the actual vs predicted plots of all the models for SPI-12.

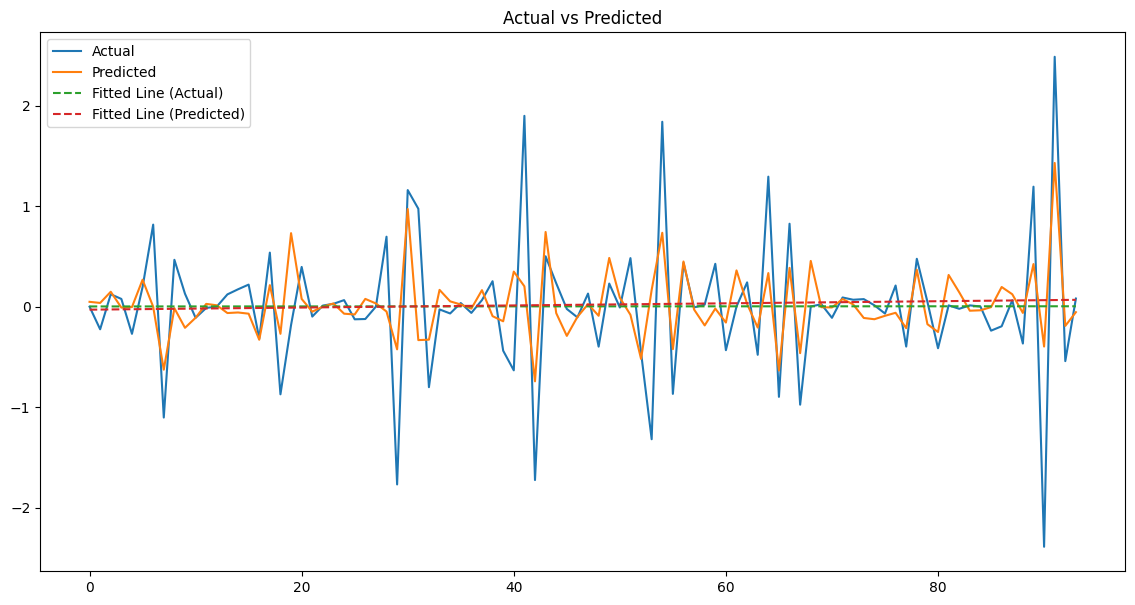
**Figure 3.6(a):** Actual Vs Predicted Plot of CNN for SPI-12



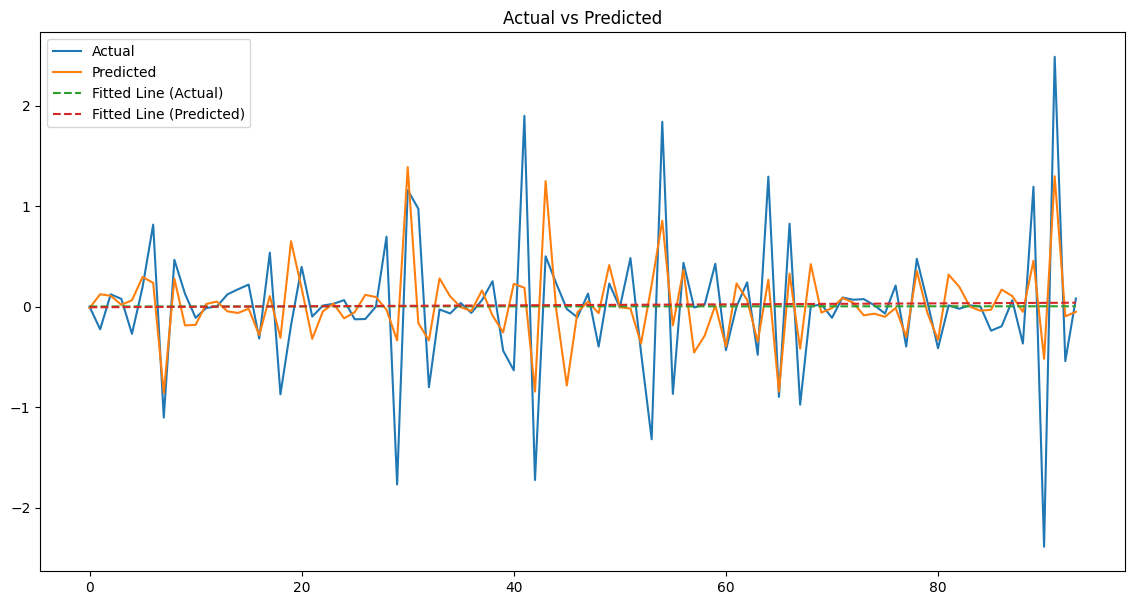
**Figure 3.6(b):** Actual Vs Predicted Plot of LSTM for SPI-12



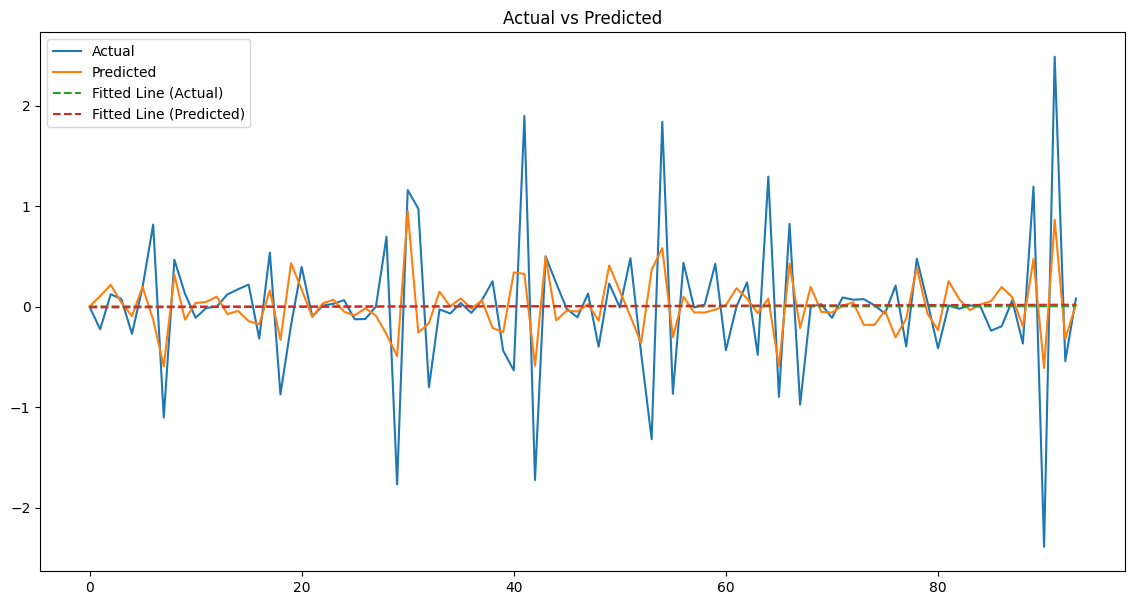
**Figure 3.6(c):** Actual Vs Predicted Plot of Adaptive Boosting for SPI-12



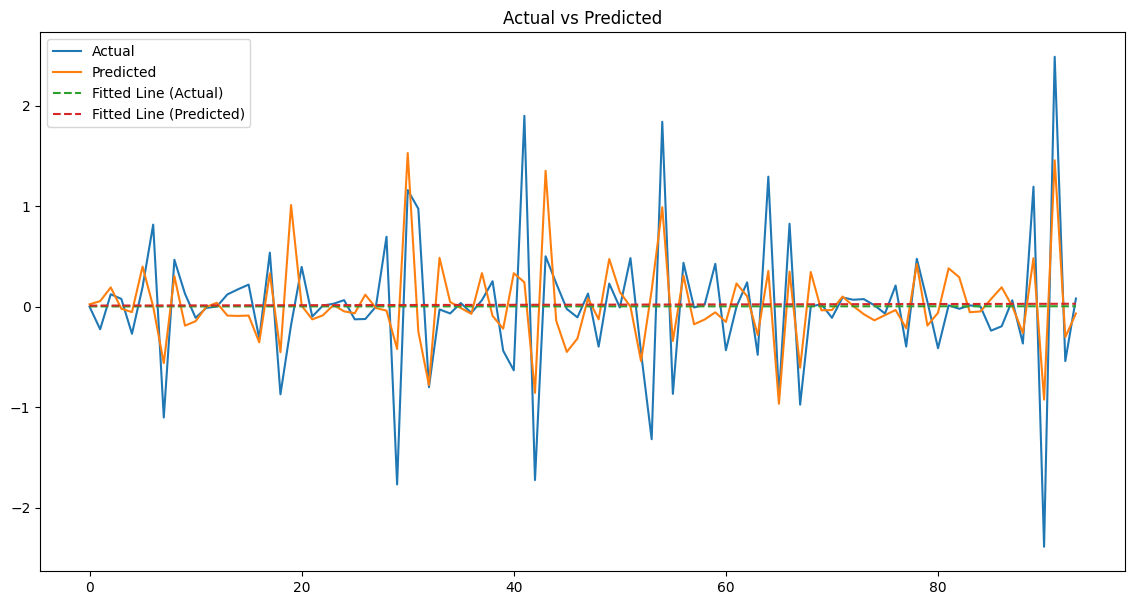
**Figure 3.6(d):** Actual Vs Predicted Plot of CatBoost for SPI-12



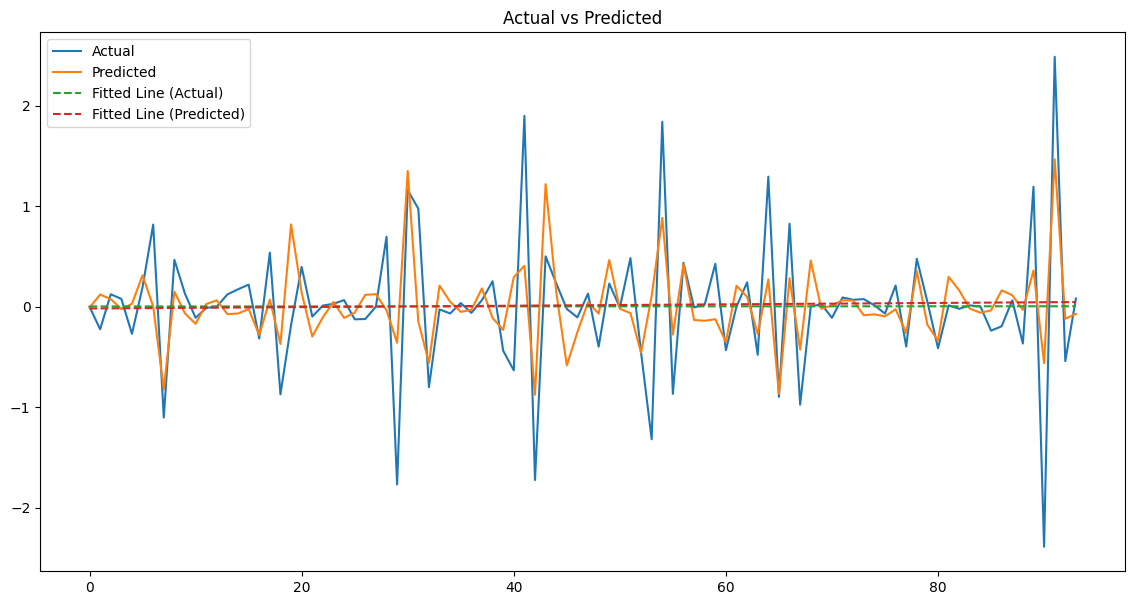
**Figure 3.6(e):** Actual Vs Predicted Plot of Gradient Boosting for SPI-12



**Figure 3.6(f):** Actual Vs Predicted Plot of KNN for SPI-12



**Figure 3.6(g):** Actual Vs Predicted Plot of Random Forest for SPI-12



**Figure 3.6(h):** Actual Vs Predicted Plot of XGBoost for SPI-12

LSTM and CNN performed the best in MAE, with the values of 31% and 32%, respectively. Random Forest, XGBoost, CatBoost and Gradient Boosting also performed decently with the values of 34% each. On the other hand, Adaptive Boosting performed the worst with the value of 40%. Similarly, in RMSE, CNN and LSTM performed the best with the values of 45% and 46%, respectively. Conversely, Adaptive Boosting Performed the worst with the value of 57%. The trend follows for NSE as well. CNN and LSTM performed the best in NSE with the values of 56% each. On the other hand, Adaptive Boosting performed the worst with the value of 30%. CNN and LSTM also performed the best in index of agreement with the values of 82% and 81%, respectively. KNN performed the worst with the value of 66%. In RAE, LSTM outperformed all other models with the lowest value of 72%. CNN was the next best model with the RAE value of 77%. Adaptive Boosting had the highest RAE value with 94%, indicating considerably worst performance. CNN and LSTM performed the best in Pearson’s Correlation Coefficient with the values of 77% each. On the other hand, Adaptive Boosting had the worst performance with the value of 56%. Table 3.3 shows the evaluation metrics of all the models for SPI-12.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Evaluation Metric | LSTM | Random Forest | KNN | Adaptive Boosting | XG Boosting | Cat Boosting | Gradient Boosting | CNN |
| MAE | 0.31 | 0.34 | 0.35 | 0.40 | 0.34 | 0.34 | 0.34 | 0.32 |
| RMSE | 0.46 | 0.50 | 0.54 | 0.57 | 0.50 | 0.52 | 0.52 | 0.45 |
| R2 | 0.56 | 0.47 | 0.37 | 0.30 | 0.46 | 0.42 | 0.43 | 0.56 |
| NSE | 0.56 | 0.47 | 0.37 | 0.30 | 0.46 | 0.42 | 0.43 | 0.56 |
| d | 0.81 | 0.77 | 0.66 | 0.68 | 0.75 | 0.71 | 0.73 | 0.82 |
| RAE | 0.72 | 0.81 | 0.83 | 0.94 | 0.80 | 0.81 | 0.81 | 0.77 |
| r | 0.77 | 0.69 | 0.66 | 0.56 | 0.69 | 0.68 | 0.66 | 0.77 |

**Table 3.3**: Evaluation Metrics of Machine Learning and Deep Learning Models of SPI-12

1. **Discussion**

Bangladesh has a mostly tropical climate with high temperatures and humidity; droughts are more common than other climatic stresses there [40]. Drought frequency, intensity, and duration vary according to hydrological, climatic, and agricultural conditions [41]. This study analyzed time-series data from meteorological stations in the Rangpur division, focusing on agricultural and hydrological drought trends. Using the SPI (Standardized Precipitation Index) indicators SPI-3, SPI-6, and SPI-12 computed from historical temperature and precipitation data, the research provides crucial understanding into drought evolution in the study region. This study represents a pioneering effort to apply deep learning techniques and machine learning techniques for drought prediction in Rangpur, Bangladesh, utilizing the SPI at three scales: SPI-3, SPEI-6, and SPEI-12. The models evaluated include K-Nearest Neighbors (KNN), Long Short-Term Memory (LSTM), Random Forest, XGBoost, AdaBoost, CatBoost, Gradient Boosting, and Convolutional Neural Networks (CNN). The study employed various performance metrics for model evaluation and comparison, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R² Score, Nash-Sutcliffe Efficiency (NSE), Relative Absolute Error (RAE), Index of Agreement (d), and Pearson's Correlation Coefficient (r).The results indicate that no specific trend was observed across the three scales SPI-3, SPI-6, and SPI-12.

The application of machine learning models ensured the robustness of drought predictions in this study. It was found that precipitation, lowest and highest temperatures, and relative humidity were critical factors for predicting droughts in Bangladesh. These findings align with those of Rahman and Lateh, who also identified similar key meteorological variables [42]. In Pakistan, significant meteorological predictors of drought included relative humidity, temperature, and wind speed [43]. These variations highlight that the importance of different meteorological parameters in drought prediction can differ based on geographic and climatic contexts.

Among the evaluated models, CNN and LSTM demonstrated the strongest performance across various metrics and time scales (SPI-3, SPI-6, and SPI-12). Specifically, CNN excelled in capturing long-term drought patterns, while LSTM showed superior performance in longer time scales (SPI-12), making it particularly effective for long-term drought prediction. XGBoost and Gradient Boosting also performed well, especially in shorter time scales, but did not consistently match the performance of CNN and LSTM across all metrics. CatBoost showed a solid performance but with slightly less consistency compared to XGBoost and Gradient Boosting. Random Forest, AdaBoost, and KNN performed less effectively in comparison to these models, with poorer R² and NSE values, particularly in capturing complex temporal patterns. These findings align with a study conducted in Turkey, which reported that CNN and LSTM-based predictive models provided superior performance in predicting drought conditions for SPI-3 and SPI-6. This consistency reinforces the effectiveness of advanced deep learning models in capturing complex temporal patterns and improving the accuracy of drought predictions across different geographic regions.

Another study conducted in Egypt revealed that CNN significantly outperformed other models in forecasting long-term drought severity. CNN’s superior performance highlighted its effectiveness and reliability for this specific application [44]. LSTM networks are specifically built to identify long-term dependencies and temporal patterns in sequential data, which makes them particularly effective for modeling long-term drought patterns. The capacity of LSTMs to retain information over long durations is well-established [45]. CNNs are known for their capacity to capture spatial hierarchies and patterns, which translate well into identifying patterns in time series data, especially at longer time scales [46]. Their superior performance in SPI-12 suggests their effectiveness in detecting extended drought conditions. Another study identified the LSTM model as the top performer for predicting both SPI and SPEI indices [47]. Besides, Random Forests and AdaBoost may not effectively capture intricate patterns due to their reliance on shallow decision trees, while KNN's performance can be hindered by high-dimensional data and lack of temporal consideration [48].

The effects of drought on agriculture in Bangladesh underscore the need for effective adaptation strategies. Changes in drought duration and frequency have led to significant crop damage and reduced cropping intensity. For instance, droughts have decreased T. Aman rice yields by 45-60% and Rabi crop yields by 50-70% [49]. To combat these issues, farmers have started using drought-resistant rice varieties, although adoption remains limited. Other strategies include altering farming practices and crop-sowing windows, which can have significant implications for food security and agricultural productivity [50-51].

Globally, the economic repercussions of severe drought vary widely based on factors like severity, economic resilience, and dependence on agriculture [52]. For example, in the United States, prolonged droughts have led to substantial economic losses, affecting both agriculture and urban water supplies [53]. Similar impacts have been observed in Australia, South Africa, India, Brazil, and countries with fragile economies like Somalia and Ethiopia [54-56]. Future research should focus on enhancing long-term drought predictions through advanced models like transformers and improved climate simulations. Such efforts will better capture climate variability and feedback mechanisms, leading to more effective drought management and resilience planning.

**Strength and limitations**

This is the first study in Bangladesh to use machine learning and deep learning models for drought prediction. Furthermore, these models were applied to predict drought using the SPI-3, SPI-6, and SPI-12 indicators. Despite advances in drought prediction models and techniques, accurate predicting drought events remains challenging due to several factors. Data availability and quality are major issues, as high-quality data on precipitation, temperature, vegetation health, climatic factors, soil moisture, and groundwater levels are essential but often limited and subject to errors, gaps, and inconsistencies. The complexity of drought processes, involving physical, climatic, social, meteorological, and economic factors, adds to the difficulty of creating accurate prediction systems. Uncertainty and variability further complicate predictions, arising from data quality, natural variability, model limitations, and climate change. The diverse temporal and spatial scales of droughts, from seasonal to decadal and local to global, present additional hurdles. Model uncertainty, influenced by the dynamic nature of drought and the selection of modeling approaches and input data, also impacts prediction accuracy. Lastly, Climate change amplifies these challenges by intensifying the frequency and severity of droughts, making it more difficult to predict such events due to the complex feedback mechanisms in changing climatic conditions.

**Recommendations**

In Bangladesh, adopting advanced predictive models such as CNN and LSTM can significantly enhance drought forecasting, given their proven capability to capture complex and long-term temporal patterns. Improving the accuracy of meteorological data collection and integrating these models into local drought monitoring systems is crucial for timely and effective drought management. Additionally, promoting drought-resistant agricultural practices and providing support to farmers can mitigate the adverse impacts on crop yields. Developing comprehensive drought management strategies and investing in climate resilience measures will further help in adapting to the changing climate. Future research should focus on incorporating advanced models and improved climate simulations to deeper understanding of and improve the management of drought risks in the region.

**Conclusion**

The study looked at the drought predictability of northwestern region of Bangladesh with machine learning and deep learning-based models. In the study, deep-learning models typically were better at predicting drought than machine learning models. In particular, CNN and LSTM performed the best in the study, scoring high in all performance metrics. Boosting algorithms such as Gradient Boosting and XGBoost generally performed well in all of SPI-3, SPI-6 and SPI-12. Conversely, KNN and Adaptive Boosting consistently performed the worst out of all models. The study suggests implementation of deep learning-based models such as CNN and LSTM for predicting drought vulnerability.

**Acknowledgement**

We are eagerly thankful to the Ministry of Science and Technology of People’s Republic of Bangladesh for their funding of this project. We also want to extend our gratitude to the Department of Environment of People’s Republic of Bangladesh for providing us the necessary data for the analysis of the project. We finally want to thank the Department of Statistics, Shahjalal University of Science and Technology, Sylhet for supporting the project all the way through.

**Conflict of Interest**

The authors declare no conflict of interest.

**Funding Information**

This study was funded by Ministry of Science and Technology of People’s Republic of Bangladesh.

**References**

1. Midgley, S.J., & Methner, N. (2016). Climate Adaptation Readiness for Agriculture: Drought Lessons from the Western Cape, South Africa.
2. Kamruzzaman, M., Almazroui, M., Salam, M.A., Mondol, M.A., Rahman, M.M., Deb, L., Kundu, P.K., Zaman, M.A., & Islam, A.R. (2022). Spatiotemporal drought analysis in Bangladesh using the standardized precipitation index (SPI) and standardized precipitation evapotranspiration index (SPEI). Scientific Reports, 12.
3. Murthy, C., Laxman, B. and Sai, M.S., 2015. Geospatial analysis of agricultural drought vulnerability using a composite index based on exposure, sensitivity and adaptive capacity. International Journal of Disaster Risk Reduction, 12: 163 -171.
4. Thomas, T., Jaiswal, R., Galkate, R., Nayak, P. and Ghosh, N., 2016. Drought indicators - based integrated assessment of drought vulnerability: a case study of Bundelkhand droughts in central India. Natural Hazards, 81(3): 1627 -1652.
5. Mohsenipour, M., Shahid, S., Chung, E. -s. and Wang, X. -j., 2018. Changing pattern of droughts during cropping seasons of Bangladesh. Water resources management, 32(5): 1555 -1568.
6. Habiba, U., Shaw, R., & Takeuchi, Y. (2012). Farmer's perception and adaptation practices to cope with drought: Perspectives from Northwestern Bangladesh. International Journal of Disaster Risk Reduction, 1, 72-84.
7. Aziz, M.A., Hossain, A., Moniruzzaman, M., Ahmed, R., Zahan, T., Azim, S., Qayum, M.A., Al Mamun, M.A., Kader, M.A., & Rahman, N.M. (2021). Mapping of Agricultural Drought in Bangladesh using Geographic Information System (GIS). Earth Systems and Environment, 6, 657 – 667
8. Fleming‐Muñoz, D. A., Whitten, S., & Bonnett, G. D. (2023). The economics of drought: A review of impacts and costs. Australian Journal of Agricultural and Resource Economics.
9. Orimoloye, I. R. (2022). Agricultural drought and its potential impacts: Enabling decision-support for food security in vulnerable regions. Frontiers in Sustainable Food Systems, 6, 838824.
10. Kamruzzaman, M., Mandal, T., Rahman, A. S., Abdul Khalek, M., Alam, G. M., & Rahman, M. S. (2021). Climate modeling, drought risk assessment and adaptation strategies in the western part of Bangladesh. Climate Vulnerability and Resilience in the Global South: Human Adaptations for Sustainable Futures, 21-54.
11. Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. Journal of hydrology, 391(1-2), 202-216.
12. Tramblay, Y., Koutroulis, A., Samaniego, L., Vicente-Serrano, S. M., Volaire, F., Boone, A., ... & Polcher, J. (2020). Challenges for drought assessment in the Mediterranean region under future climate scenarios. Earth-Science Reviews, 210, 103348.
13. Dikshit, A., Pradhan, B., & Santosh, M. (2022). Artificial neural networks in drought prediction in the 21st century–A scientometric analysis. Applied So Computing, 114, 108080.
14. Bochenek, B., & Ustrnul, Z. (2022). Machine learning in weather prediction and climate analyses—applications and perspectives. Atmosphere, 13(2), 180.
15. Lee, D., Kang, S., & Shin, J. (2017). Using deep learning techniques to forecast environmental consumption level. Sustainability, 9(10), 1894.
16. Feng, P., Wang, B., Li Liu, D., & Yu, Q. (2019). Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. Agricultural Systems, 173, 303-316.
17. Gyaneshwar, A., Mishra, A., Chadha, U., Raj Vincent, P. D., Rajinikanth, V., Pattukandan Ganapathy, G., & Srinivasan, K. (2023). A Contemporary Review on Deep Learning Models for Drought Prediction. Sustainability, 15(7), 6160.
18. Danandeh Mehr, A., Rikhtehgar Ghiasi, A., Yaseen, Z. M., Sorman, A. U., & Abualigah, L. (2022). A novel intelligent deep learning predictive model for meteorological drought forecasting. Journal of Ambient Intelligence and Humanized Computing, 1-15.
19. Shahdad, M., & Saber, B. (2022). Drought forecasting using new advanced ensemble-based models of reduced error pruning tree. Acta Geophysica, 70(2), 697-712.
20. Abdar, M., Pourpanah, F., Hussain, S., Rezazadegan, D., Liu, L., Ghavamzadeh, M., ... & Nahavandi, S. (2021). A review of uncertainty quantification in deep learning: Techniques, applications and challenges. Information fusion, 76, 243-297
21. Shahid, S. & Behrawan, H. Drought risk assessment in the western part of Bangladesh. Nat. Hazards https://doi.org/10.1007/s11069-007-9191-5 (2008).
22. Keyantash, J. & Dracup, J. A. Te quantifcation of drought: An evaluation of drought indices. Bull. Am. Meteorol. Soc. 83, 1167–1180 (2002).
23. Mundial, B. Water resource management in Bangladesh: Steps towards a new national water plan. In Water resource management in Bangladesh: Steps Towards a new national water plan (Banco Mundial, 1998).
24. Abdullah, H. M., & Rahman, M. M. (2015). Initiating rain water harvest technology for climate change induced drought resilient agriculture: scopes and challenges in Bangladesh. Journal of Agriculture and Environment for International Development (JAEID), 109(2), 189–208. <https://doi.org/10.12895/jaeid.20152.334>
25. Mirza, M.Q. and Paul, S., (1992) Prakritik Durgojob O Bangladesh Paribesh (Natural Disaster and Environment in Bangladesh), Centre for Environmental Studies and Research, Dhaka.
26. Rahman, M. S., & Islam, A. R. M. T. (2019). Are precipitation concentration and intensity changing in Bangladesh overtimes? Analysis of the possible causes of changes in precipitation systems. *Science of the Total Environment*, *690*, 370-387.
27. Farahmand, A., & AghaKouchak, A. (2015). A generalized framework for deriving nonparametric standardized drought indicators. *Advances in Water Resources*, *76*, 140-145.
28. Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*, 197-227.
29. Wu, J. (2017). Introduction to convolutional neural networks. *National Key Lab for Novel Software Technology. Nanjing University. China*, *5*(23), 495.
30. Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia*, *4*(2), 1883.
31. Schapire, R. E. (2013). Explaining adaboost. In *Empirical inference: festschrift in honor of vladimir N. Vapnik* (pp. 37-52). Berlin, Heidelberg: Springer Berlin Heidelberg.
32. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
33. Hancock, J. T., & Khoshgoftaar, T. M. (2020). CatBoost for big data: an interdisciplinary review. *Journal of big data*, *7*(1), 94.
34. Li, C. (2016). A gentle introduction to gradient boosting. *URL: http://www. ccs. neu. edu/home/vip/teach/MLcourse/4\_ boosting/slides/gradient\_boosting. pdf*, *59*.
35. Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270.
36. Prasad, R., Deo, R. C., Li, Y., & Maraseni, T. (2018). Ensemble committee-based data intelligent approach for generating soil moisture forecasts with multivariate hydro-meteorological predictors. *Soil and Tillage Research*, *181*, 63-81.
37. Choesang, T., Ryntathiang, S., Jacob, B. A., Krishnan, B., & Kokatnoor, S. A. (2022, December). Drought Prediction—A Comparative Analysis of Supervised Machine Learning Techniques. In *Congress on Smart Computing Technologies* (pp. 295-307). Singapore: Springer Nature Singapore.
38. Hukkeri, G. S., Naganna, S. R., Pruthviraja, D., Bhat, N., & Goudar, R. H. (2023). Drought forecasting: Application of ensemble and advanced machine learning approaches. *IEEE Access*.
39. Jiang, W., & Luo, J. (2022). An evaluation of machine learning and deep learning models for drought prediction using weather data. *Journal of Intelligent & Fuzzy Systems*, *43*(3), 3611-3626.
40. Yaseen, Z. M., Ali, M., Sharafati, A., Al-Ansari, N. & Shahid, S. Forecasting standardized precipitation index using data intelligence models: Regional investigation of Bangladesh. *Sci. Rep.* **11**, 1–25 (2021).
41. Heim, R. R. A review of twentieth-century drought indices used in the United States. *Bull. Am. Meteorol. Soc.* **83**, 31 (2002).
42. Rahman, M. R. & Lateh, H. Meteorological drought in Bangladesh: Assessing, analysing and hazard mapping using SPI, GIS and monthly rainfall data. *Environ. Earth Sci.* **75**, 22 (2016).
43. Khan, N. *et al.* Prediction of droughts over Pakistan using machine learning algorithms. *Adv. Water Resour.* **139**, 103562 (2020).
44. Elbeltagi A, Srivastava A, Ehsan M, Sharma G, Yu J, Khadke L, Gautam VK, Awad A, Jinsong D. Advanced stacked integration method for forecasting long-term drought severity: CNN with machine learning models. Journal of Hydrology: Regional Studies. 2024 Jun 1;53:101759.
45. Olah, C. (2015). Understanding lstm networks.
46. LeCun Y, Bengio Y, Hinton G. Deep learning. nature. 2015 May 28;521(7553):436-44.
47. Poornima S, Pushpalatha M. Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. Soft Computing. 2019 Sep 1;23:8399-412.
48. Breiman L. Random forests. Machine learning. 2001 Oct;45:5-32.
49. Habiba, U., Shaw, R. & Takeuchi, Y. Drought risk reduction through a socio-economic, institutional and physical approach in the northwestern region of Bangladesh. *Environ. Hazards* **10**, 31 (2011).
50. Nayak, S. *et al.* Adoption trend of climate-resilient rice varieties in Bangladesh. *Sustainability* **14**, 5156 (2022).
51. Ahmed, B. *et al.* Indigenous people’s responses to drought in northwest Bangladesh. *Environ. Dev.* **29**, 4 (2019).
52. Shayanmehr, S. *et al.* The impacts of climate change on water resources and crop production in an arid region. *Agriculture* **12**, 1056 (2022).
53. Easterling, D. R. *et al.* Climate extremes: observations, modeling, and impacts. *Science (80-)* **289**, 2068–2074 (2000).
54. Choruma, D. J., Akamagwuna, F. C. & Odume, N. O. Simulating the impacts of climate change on maize yields using EPIC: A case study in the Eastern Cape Province of South Africa. *Agriculture* **12**, 794 (2022).
55. Auffhammer, M., Hsiang, S. M., Schlenker, W. & Sobel, A. Using weather data and climate model output in economic analyses of climate change. *Rev. Environ. Econ. Policy* (2013).
56. Assunção, J. & Chein, F. Climate change and agricultural productivity in Brazil: Future perspectives. *Environ. Dev. Econ.* **21**, 581–602 (2016).