TFT Architecture and RNN Variants for Water Quality Prediction of Bharathapuzha River

Jitha P Nair[1*], Vijaya M S[2]

[1*]Research Scholar, Department of Computer Science, PSGR Krishnammal College for Women, Peelamedu, Coimbatore, India
[2]Associate Professor, Department of Computer Science, PSGR Krishnammal College for Women, Peelamedu, Coimbatore, India

Abstract : Water quality is a crucial aspect of the health and well-being of communities and ecosystems around the world. This study focuses on water quality prediction for the Bharathapuzha River, which is susceptible to various sources of contamination. The prediction model is built using Recurrent Neural Network (RNN) variants and Temporal Fusion Transformers (TFT) approach. The dataset developed for building the prediction model consists of 2190 unique instances containing physicochemical and seasonal parameters. Water quality varies over time due to changes in natural and human factors, such as weather conditions, land use, pollution levels, and treatment processes. The objective is to capture the time series patterns in the data and to forecast the water quality index accurately. The performance result demonstrates that the TFT outperforms the RNN variants in prediction. This study highlights the importance of TFT in trend analysis and developing a reliable forecasting model.

Keywords: River Water Quality, Prediction Model, Deep Learning Architectures, Physicochemical Parameters, Seasonal Parameters

1. INTRODUCTION

The precise forecasting of water quality is imperative for developing effective management approaches and is thus a primary concern for water resource managers and policymakers in ensuring environmental sustainability and public health. The Bharathapuzha River is a vital water resource in the Indian state of Kerala but is increasingly threatened by pollution from various sources, including agricultural runoff, industrial discharges, and municipal waste. Predicting water quality in the river is a challenging task, given the complexity of the underlying processes and the interplay of multiple environmental and physicochemical factors.

In recent years, machine learning algorithms have emerged as a promising approach to predicting water quality based on a range of factors. These algorithms have the potential to capture complex patterns in large datasets and improve the accuracy of water quality predictions. However, there is a need for more research on the use of advanced machine learning techniques, such as deep learning, in predicting water quality for complex river systems such as the Bharathapuzha River.

In this research article, proposes a water quality prediction model for the Bharathapuzha River that employs Recurrent Neural Network (RNN) variants and Temporal Fusion Transformers (TFT). The model aims to capture the physicochemical and seasonal patterns in the time series dataset to forecast water quality accurately. The dataset with 2190 unique instances that contain both physicochemical and seasonal parameters to build the prediction model. The use of RNN variants and TFT allows us to capture long-term temporal dependencies and improve the accuracy of the prediction. The study contributes to the growing body of research on the use of deep learning algorithms in predicting water quality for complex river systems. The proposed model has the potential to help water resource managers in developing effective strategies for water quality management and conservation. Additionally, the study highlights the importance of advanced machine learning techniques for predicting water quality in complex river systems and provides insights into the potential of deep learning for addressing water quality challenges.

Numerous research studies have been conducted utilising a restricted set of physicochemical parameters to construct water quality, forecasting models. However, it has been observed that enhancing the number of physicochemical factors, in addition to incorporating seasonal variables, can effectively enhance the efficiency of water quality prediction.

The grey relational method, mathematical statistics method, model-based approach, Bayesian approach, Genetic Algorithm, MLP regressor, and support vector regressor are computational methods used by researchers currently in the existing water quality prediction research.

Umair Ahmed et al. [2] established an efficient water quality prediction Framework with supervised machine learning. This framework provides a strategy that employs four information boundaries, namely temperature, turbidity, pH,

_____

and solids that have been entirely dispersed. The research was supported by data from PCRWR, which included 663 samples from 12 distinct wellsprings in Pakistan's Rawal Lake. WQI was evaluated utilising a range of AI calculations directed by agents and also calculating relapse and categorization. The eight lapse calculations for WQI and 10 classification calculations for ordering experiments into predetermined WQC computations had been assessed.

Shuangyin Liu et al. [5] proposed a half-and-half methodology of help vector relapse with hereditary calculation advancement for hydroponics water quality prediction. This paper proposes a forecast model based on help vector relapse (SVR) to address the hydroponics water quality expectation issue. When putting together a successful SVR model, the SVR boundaries should be set with caution. This study presents a half-and-half methodology known as genuine worth hereditary calculation uphold vector regression (RGA-SVR), which looks for the best SVR boundaries using genuine esteemed hereditary calculations and then uses the best boundaries to build the SVR models. The methodology is used to forecast the hydroponics water quality data collected from Yixing's oceanic plants in China. The results show that RGA-SVR outperforms the standard SVR and back-engendering (BP) neural organisation models based on the root mean square error (RMSE) and mean outright rate blunder (MAPE). This RGA-SVR model is a viable method of dealing with anticipated hydroponics water quality.

Salisu Yusuf Muhammad et al. [6] introduced a machine learning-based water quality classification model. This article proposes a reasonable grouping model for ranking water quality based on AI calculations. The paper separated and analysed the presentation of various arrangement models and calculations to identify the key factors that contributed to the classification of the water nature of the Kinta River in Perak, Malaysia. Five mathematically precise models were evaluated, compared, and displayed. The Lazy model utilising the K Star calculation was the most accurate grouping model among the five models, with an exactness of 86.67%. In general, wastewater is hazardous to human health, and it is essential to develop logical models to address this problem.

Liao and Zhao [18] focused on dissolved oxygen for water quality prediction and proposed a combined model consisting of fuzzy neural networks (FNN), principal component analysis (PCA), and differential evolution by the BP algorithm (DEBP). PCA contributes to the dimension reduction of the input data vector and differential evolution algorithm.

Wang et al. [17] demonstrated an LSTM (long- and short-term memory) neural network-based deep learning approach.

The LSTM NN model was constructed for prediction, followed by the collection of training data from Taihu Lake and the selection of appropriate parameters to improve neural network accuracy. Due to the nonlinear, dynamic, changing, and complex nature of the water parameter quality parameters, predicting WQ is a hard task. Due to these traits, traditional forecasting algorithms suffer from poor accuracy and increased processing complexity.

From the existing research in water quality index prediction, several limitations have been identified. These include insufficient data, inadequate coverage, complex parameters, limited indices, and lack of standardization. Insufficient data and inadequate coverage of monitoring stations in developing countries limit the accuracy and reliability of water quality index predictions. The complex interrelatedness of water quality parameters makes it challenging to model and predict their interactions accurately, while the use of a limited number of water quality indices can result in oversimplified predictions. Additionally, the absence of standardization in water quality index calculations and interpretation can lead to inconsistent and unreliable results. Therefore, addressing these limitations and developing more accurate and reliable water quality index prediction models can be a crucial step towards improving water management and ensuring safe water for various uses.

The objective of this study is to develop an improved water quality prediction model through the utilization of recurrent neural network variants and a temporal fusion transformer for Bharathapuzha river water data. In order to achieve this objective, physiochemical time series data from the Bharathapuzha River is collected from three sampling stations and the seasonal parameters are collected from the visual crossing site based on the location of sampling stations. The proposed model for river water quality prediction employs various deep learning architecture RNN variants such as Long Short-Term Memory, Gated Recurrent Unit, and Temporal Fusion Transformer. The performance of these models is assessed to determine their effectiveness in enhancing the prediction accuracy of water quality.

2. DATA COLLECTION AND DATASET PREPARATION

In this research, a water quality predictive model is constructed by identifying the trends from physicochemical and seasonal features in time series river water data. The raw data with 26 physicochemical parameters such as pH, conductivity, turbidity, phenolphthalein alkalinity, total alkalinity, chloride, COD, TKN, ammonia, Ca. hardness, Mg. Hardness, sulphate, sodium, TSS, TDS, FDS, phosphate, boron, potassium, BOD, fluoride, Nitrate-N, DO,

823

_____

TC and FC is collected from three sampling stations of Bharathapuzha River. The seasonal characteristics are obtained from the visual crossing website corresponding to locations of the sampling station across the river between January 2020 and December 2021. The analysis focuses on seasonal parameters, including dew, humidity, sea level pressure, precipitation, precip over, wind speed, wind direction, cloud cover, and visibility, which exhibit significant alterations across different seasons. These seasonal attributes are fused with physicochemical parameters in the development of a comprehensive dataset for this study.

Physiochemical Parameters

Physicochemical parameters are essential indicators of water quality, helping assess its suitability for different purposes. The 26 commonly used parameters include pH, which measures water acidity or alkalinity, and conductivity, reflecting the presence of dissolved salts and minerals. Turbidity indicates the water's cloudiness or haziness caused by suspended particles. The phenolphthalein and total alkalinity measure the ability of water to neutralize acids. Chloride and COD are used to assess contamination and pollution levels, while TKN and ammonia indicate nitrogen content. Calcium and magnesium hardness, sulphate, and sodium can contribute to scaling, corrosion, and other issues. TSS, TDS, FDS, phosphate, boron, potassium, BOD, fluoride, and nitrate-n affect the taste of water, odour, clarity, and nutrient content. DO is critical for aquatic life, while TC, FC, and total coliforms indicate the presence of pathogens and faecal contamination.

Seasonal Parameters

The impact of seasonal parameters on river water quality is influenced by sudden climate changes. Literature indicates that seasonal parameters influence the water quality index and its prediction over time series data. The relationship between simultaneous rainfall and humidity is strong, with relative humidity improving due to the evaporation of rainwater. The Davis Cup Anemometer is utilized to measure wind speed at a 3-meter height, providing a comparison with traditional 10-meter measurements. An increase in wind speed results in decreased transition time between evaporative stages at low-velocity values. Dew is a crucial source of river water, significantly impacting microclimates and the physiological state of vegetation. Temperature plays a crucial role in this assessment, serving as an indicator of certain species and the water body's activity. Humidity, another measure of atmospheric water vapor, holds similar importance in measuring water quality due to its potential impact on evaporation rates and temperature of the environment. Global warming will alter precipitation distribution by altering air temperatures and circulation patterns, further exacerbating the impact of seasonal factors on water quality. The alteration of physicochemical parameter acceptable limits due to seasonal factors results in lower water quality. This work considers seasonal features for the same study period, emphasizing their significance in improving the efficiency of the predictive models.

WQI Calculation

The WQI serves as a comprehensive measure of the water quality for the proposed system. It facilitates the monitoring of water quality changes over time and provides an assessment of the suitability of the water body The WQI is determined by the average of several indicators such as dissolved oxygen, pH, nutrient levels, and turbidity. It is then assigned a score on a scale ranging from 0 to 120, with higher scores indicating poor water quality. In this regard, the WQI is computed and included as the target variable along with 40 independent variables for the WQI modelling prediction task. The dataset incorporates both physiochemical and seasonal parameters and comprises 2190 instances.

The present study relied on a comprehensive dataset comprising twenty-six physiochemical attributes, ten seasonal attributes, and spatial parameters, referred to as the WQI-BP and tabulated in Table 1. The resulting time series data encompassed a total of 40 attributes, meticulously prepared to build the WQI prediction models.

Table 1: Sample Water Quality Data Collected from Sampling Stations and Visual Crossing Site

Temp	29	27	30	28	30	30	30	28	29	29	29	29	27	28	28	30	28	27	30
pH	7.31	7.60	7.43	7.27	7.63	7.20	7.52	7.47	7.38	7.34	7.36	7.28	7.32	7.21	7.61	7.68	7.34	7.20	7.61
Conductivity	312	316	295	290	316	293	295	308	312	293	303	306	301	310	304	293	303	300	293

Turbidity	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
PhenolphthAlkalinity	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total Alkalinity	77	75	76	79	80	78	79	76	79	80	80	78	75	80	78	78	80	76	75
Cholride	45	43	39	38	39	39	42	41	43	43	40	44	43	40	42	43	43	38	44
COD	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
TKN	0.1	0.1	0.09	0.1	0.1	0.09	0.1	0.1	0.1	0.11	0.11	0.1	0.1	0.1	0.1	0.1	0.11	0.11	0.1
Ammonia	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Hardness	105	100	104	108	103	107	105	105	108	107	106	101	107	100	101	107	103	108	104
Ca.Hardness	54	50	51	53	50	50	51	53	51	50	52	50	54	53	51	52	52	52	51
Mg.Hardness	53	51	50	50	52	50	52	52	52	53	50	53	51	50	54	52	52	52	53
Sulphate	0.10	0.31	0.14	0.28	0.31	0.18	0.07	0.21	0.58	0.60	0.59	0.68	0.61	0.79	0.51	0.54	0.59	0.71	0.71
Sodium	16	15	17	17	18	15	15	15	17	18	15	16	16	17	17	15	15	15	15
TDS	180	173	171	175	180	177	180	172	182	179	181	178	173	171	174	171	171	175	180
FDS	99	101	94	102	101	100	100	95	94	102	96	97	101	99	95	98	101	94	94
TSS	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300

Phosphate	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.006	0.006
Boron	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Pottassium	6.38	6.40	6.46	6.42	6.45	6.41	4.10	4.25	4.45	4.21	4.49	3.41	3.34	3.43	3.18	3.11	3.15	3.44	3.24
BOD	1.01	1.03	1.02	1.02	1.03	1.02	1.00	1.03	1.00	1.00	1.01	1.00	1.01	1.01	1.01	1.01	1.20	1.24	1.05
Fluoride	0.58	0.45	0.50	0.45	0.46	0.56	0.43	0.37	0.41	0.33	0.53	0.10	0.28	0.40	0.11	0.22	0.16	0.10	0.11
Nitrate-N	1.29	1.03	1.51	1.19	1.78	1.49	0.92	1.29	1.10	1.76	0.41	0.91	0.61	0.50	0.31	1.15	0.93	0.96	0.76
TC	320	306	338	312	261	338	323	293	313	311	278	334	280	326	324	253	300	268	252
FC	281	272	258	264	254	296	265	254	265	260	299	260	259	298	265	282	251	281	253
DO	6.12	6.25	6.12	6.25	6.24	6.18	6.26	6.06	6.06	6.07	6.28	6.23	6.03	6.04	6.27	6.22	6.02	6.23	6.11
Dew	15.7	14.6	13.4	13.6	15.6	17.7	18.9	19.4	18.3	17.8	17.5	17.3	17.5	17.4	17.3	18.3	18.7	19	20.1
Humidity	59.30	56.72	51.89	53.06	58.80	62.79	68.91	68.63	65.71	63.80	63.33	63.55	65.59	64.16	62.20	65.17	67.23	65.80	69.30
Sealevelpressure	1016.60	1017.10	1015.80	1015.70	1014.80	1014.80	1015.50	1015.50	1013.70	1014.50	1014.20	1013.40	1013.20	1012.90	1012.70	1011.90	1011.80	1010.90	1010.80
Precipitation	0	2	0	4	18	19.41	14.81	14	5	11.85	13.6	4	35.17	9.56	17.7	14.55	2.06	0.5	24.5

Precipc over	0	4.17	0	8.33	4.17	8.33	8.33	4.17	8.33	8.33	4.17	4.17	12.5	4.17	12.5	12.5	12.5	4.17	4.17
Windspeed	16.3	14.4	13.1	15.4	14	18.7	40.2	13.6	14.4	14.9		14	13.7	13.3	11.9	9.4	11.6	11.6	9.9
Winddir	52.9	62.3	61.7	68.2	56.5	69.3	114.6	95	94.9	65.1	99.1	96.3	84.6	51.8	72.7	140.8	100.4	71.9	129.3
Cloud cover	27.4	17.9	5.5	14.1	14.6	16	32.3	42.5	26.3	14	14.3	12.5	30.8	22.9	23.2	26.1	27.4	33.1	71.1
Visibility	5.5	6	5.7	5.9	5.6	5.5	4.8	5.3	5.1	5.4	5.5	5.7	5.2	5.2	5.4	5.2	5.3	5.6	5.2

Exploratory Data Analysis

The collected river water quality data is subjected to Exploratory Data Analysis (EDA) in order to comprehend the properties of the data and evaluate the significance of each parameter in generating the water quality index. Statistical methods, such as heatmaps, boxplot analysis, pair plot analysis, and histograms are utilized to analyse and understand the distribution of parameter values. Fig. 1a depicts a bar graph analysis of humidity, wind speed, cloud cover, visibility, and physicochemical parameters while Fig.1b displays the parameters that negatively correlated with WQI, including pH, turbidity, FDS, TSS, boron, TC, cloud cover, and wind speed. Boxplot studies revealed that seasonal parameters like wind speed and cloud cover exhibited a wide range of values, varying between 10 and 270 for wind speed and 0 to 100 for cloud cover as illustrated in Fig.1c. Min-Max approach is used to normalise parameter values to standard ranges, wind speed and cloud cover are standardised. Additionally, a heatmap is used alongside Pearson correlation to visualize and analyse the correlation, both positive and negative, between parameters.
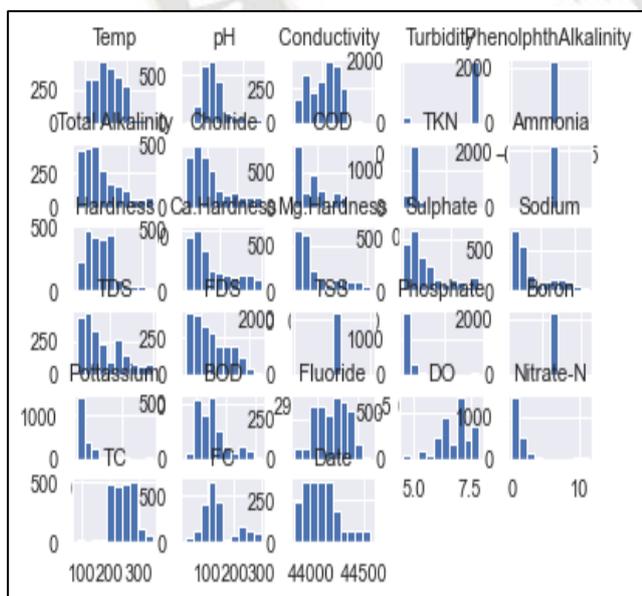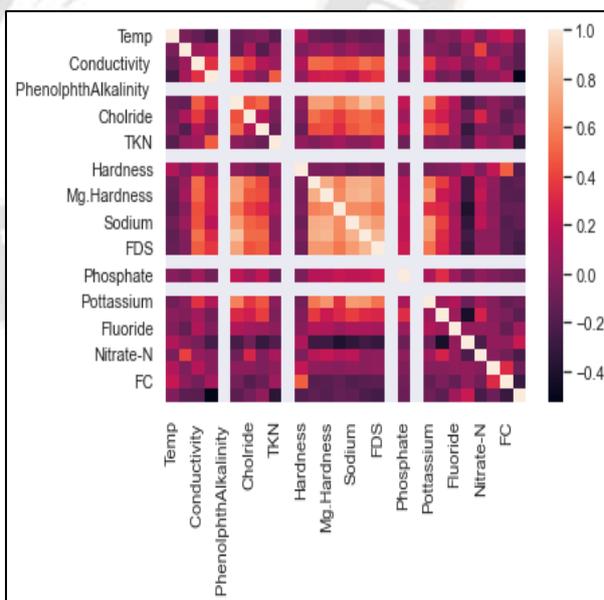

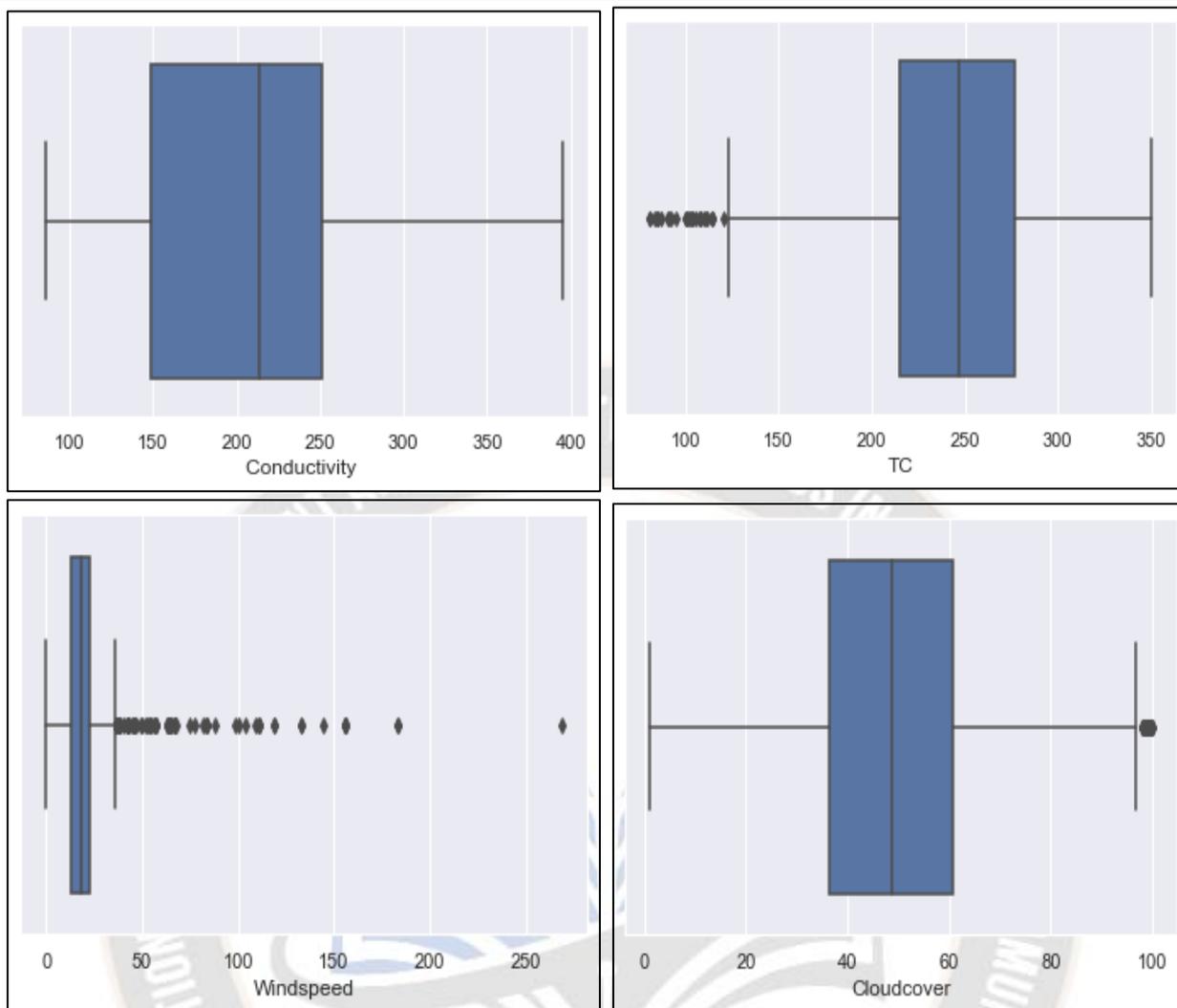Fig.1a Bar graph analysis

Fig.1b Heatmap analysis

Fig. 1c. Boxplot Analysis of Attributes

Through the process of EDA, it is discovered that certain instances within the dataset contained missing values which required elimination. Consequently, data cleaning is performed to ensure data accuracy. The min-max normalization is applied to parameters such as conductivity, total coliform, wind speed and cloud cover to be in a normal range. The EDA provided valuable insights into attribute distributions and parameter correlations, thereby presenting appropriate solutions for data modelling and pre-processing requirements.

Data Pre-processing

Feature selection is a crucial process in predictive modelling, involving the identification of relevant parameters that significantly contribute to predicting the target variable. In this study, the select K best algorithm was utilized to determine essential features in calculating the water quality index. Based on the select K best feature selection algorithm, conductivity ranked first in estimating the water quality index, followed by ammonia and phosphate, while the

negatively ranked attributes such as boron and phenolpth alkalinity were deemed insignificant and, thus, removed from the dataset. The feature selection method enhanced the river water quality dataset, resulting in the development of the WQI-BP dataset, which comprises 2190 instances and 38 attributes and is a valuable resource for training the deep neural network.

3. WQI PREDICTION MODEL

The present study offers a solution to the challenge of anticipating the water quality index by framing it as a regression problem tackled through RNN variants and TFT architecture. RNN variants are highly adept at characterizing, classifying, and effectively modelling data inputs, thanks to their intricate connection of nodes with multiple hidden layers, where two visible layers are deployed as input and output layers to strengthen the predictive accuracy of the model. TFT works by combining multiple layers of transformer blocks to capture long-term dependencies in the time series data. The input layer takes in the input WQI-BP

time series data and converts it into a numerical representation. The encoder layers use self-attention mechanisms to process the input data and generate feature representations for each instance in the series and allow information from past and future time steps to be fused into the representation for each step. The architecture of the proposed WQI prediction model is illustrated in Fig. 3, where the pre-processed data is fed through the input layer of the RNN variants and TFT architecture and the final prediction is made through its output layer.
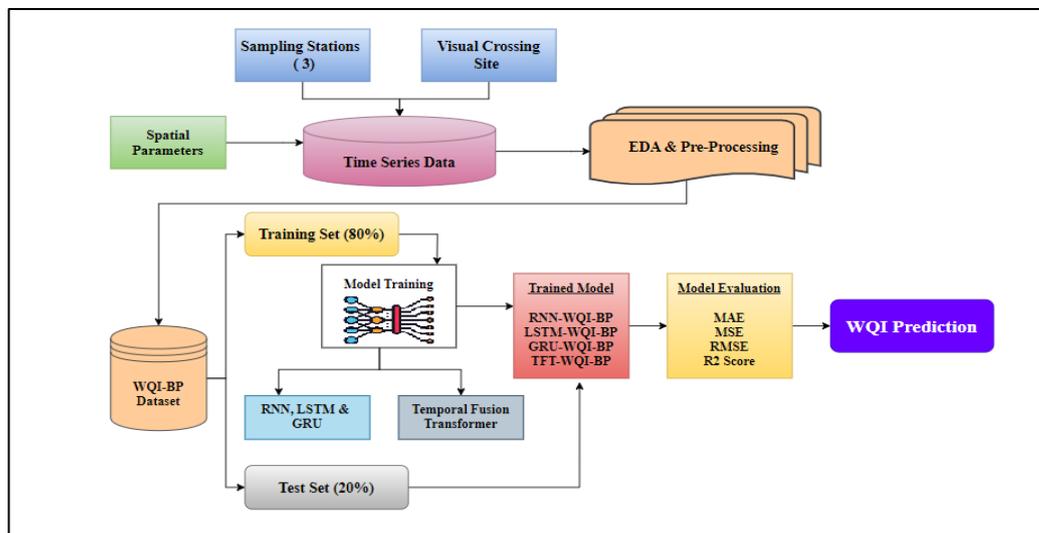
Fig. 2   Proposed WQI Model Architecture

The utilization of deep learning architectures, including recurrent neural networks, long short-term memory, and gated recurrent networks, has been specifically designed and developed to facilitate the training of sequence data. These architectures have been chosen in this research work to construct a river water quality index prediction model. Among these architectures, Recurrent Neural Network (RNN) stands out as it utilizes the result from the preceding section as input for the next. The Hidden state, which stores crucial information about a sequence, is the primary component of RNN. However, RNNs are susceptible to the vanishing gradients problem due to their limited ability for long-term memory. The primary challenge for RNN is maintaining data consistency across several time steps. To address this issue, gated recurrent networks and Long Short-term Memory have been employed. Long Short-term Memory (LSTM) recurrent unit aims to recall all the earlier data encountered by the network and to forget irrelevant information. Furthermore, each LSTM recurrent unit stores a vector referred to as the Internal Cell State, which conceptually describes the information retained by the previous LSTM recurrent unit. On the other hand, Gated Recurrent Unit (GRU) tackles the vanishing gradient problem by utilizing an update gate and reset gate, acting as two vectors that determine the information sent to the output. One distinguishing feature of these deep learning architectures is that they are trained to retain knowledge from

a long time ago without erasing it or removing extraneous data.

Temporal Fusion Transformer

The Temporal Fusion Transformer is a deep learning model that is proposed for the prediction of time-series WQI data. TFT is a fusion of various neural networks, including feedforward neural networks, convolutional neural networks, and the Transformer. The model uses a hierarchical structure to capture different levels of temporal dependencies and combines them to make accurate predictions. TFT can handle complex and dynamic patterns in time-series data, making it suitable for a wide range of applications such as traffic forecasting, energy demand prediction, and financial time series analysis.

The importance of TFT in time series data analysis lies in its ability to handle complex and high-dimensional data. Unlike traditional time series models, such as ARIMA and exponential smoothing, TFT can handle multiple input time series with different temporal resolutions, missing data, and irregularly sampled data. TFT also incorporates attention mechanisms to enable the model to focus on specific parts of the input time series, allowing it to capture complex and nonlinear relationships between different variables. Furthermore, TFT can handle both continuous and categorical features, making it versatile for a wide range of time series forecasting tasks.

The benefits of using TFT include its ability to handle a wide range of temporal patterns, including seasonality, long-term

829

_____

dependencies, and irregularities in the data. TFT is also designed to handle missing data and can effectively handle situations where data is irregularly sampled. Another benefit of TFT is that it can provide interpretability, allowing users to understand the factors driving the predictions. TFT has shown promising results in various real-world applications, including energy demand prediction and stock price forecasting, making it a valuable tool for time-series analysis.

The Temporal Fusion Transformer is an architecture designed for time series forecasting tasks, illustrated in Fig. 3. It incorporates the standard Transformer architecture with the idea of temporal convolution and gating mechanisms. TFT has an encoder-decoder structure, where the encoder is responsible for capturing the input time series temporal dependencies and the decoder is responsible for generating the output forecasts. TFT uses a temporal convolution layer in the encoder to capture local temporal dependencies. It also uses a gating mechanism to selectively weigh the importance of past and present information for predicting the future. Additionally, TFT uses an autoregressive approach to generate the output forecasts. The decoder takes the encoder's output and recursively generates future time steps one by one. TFT also includes a multi-scale input module that allows the model to capture patterns at different time scales. It also has a hybrid attention mechanism that incorporates both local and global temporal dependencies.
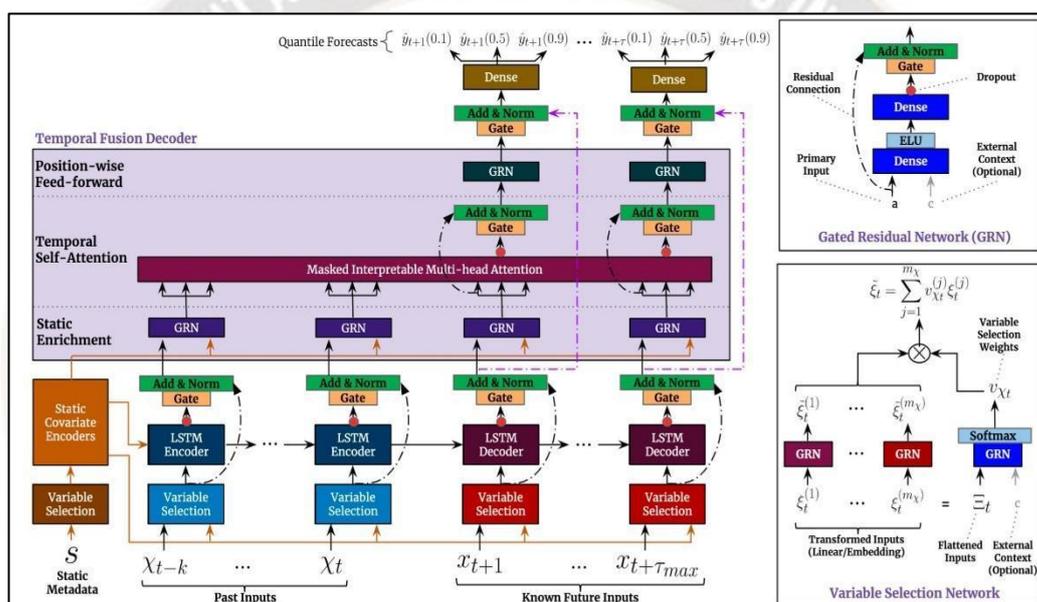

Fig.3. Architecture of Temporal Fusion Transformer

Model building and Evaluation

The 80% of instances of the WQI-BP dataset prepared as above are given as input to RNN and its variants LSTM, GRU and TFT architecture for training the networks. The best hyperparameters are chosen during model training to make the model more effective at mapping the input features as independent variables to the target variable as the dependent variable.

Hidden layers, dense layers, optimizer, epoch, momentum, batch size, activation function, and dropout are examples of hyperparameters that are utilised in deep learning architectures to enhance model accuracy and fine-tune the WQI forecasting model. Hidden layers are the layers that are in between the input and output layers. A layer that is densely connected is one in which each layer receives input from all of the layers below it. The range is set between 5 and 10 units, and dense layers improve overall accuracy. Optimizers are methods that alter the properties of the neural network, like its weights and learning rate, to reduce losses and address optimization issues. The number of dataset complete iterations required is determined by the epoch size. Momentum is a unique hyperparameter that enables the search direction to be determined not only by the gradient from the current step but also by the gradients from previous steps. The model's nonlinearity is introduced through activation functions. The activation function can split them into different layers and get a reduced output of the density layer. By passing randomly selected layers and limiting sensitivity to particular layer weights, the dropout layer helps prevent training overfitting. The speed at which a deep model replaces a previously learned concept with a new one is determined by the learning rate. Finally, the WQI prediction models are built by representation learning from the input instances using RNN, LSTM, GRU and TFT with proper

_____

hyperparameters settings.

The effectiveness of the WQI forecasting model is evaluated using the evaluation metrics such as R2 score, root mean squared error, mean squared error, and mean absolute error. An estimator's mean squared error is the average of error squares or the difference between the predicted value and the actual value. The average difference between the predicted value and the actual value is what is used to calculate the mean absolute error. The root mean square error is used to measure a model's prediction error for quantitative data, which is a metric that indicates how well a regression line fits the data points. The R2 score value determines the accuracy of the model. If the R2 score value is high then the model is considered to be good in predicting the target variable and if the R2 score is less than 0.5 then the model is not considered to be good. The prediction models are found to be effective when the error rate is less with a high R2 score value. In this work, the performance of the WQI predictive models built with physicochemical and seasonal data is evaluated using the metrics with 20% of the dataset as the test set.

4. EXPERIMENT AND RESULTS

In this work, using Python libraries, the experiments were carried out by implementing the deep learning architectures. The RNN, LSTM, GRU and TFT networks have been trained with the training dataset WQI-BP, which contains tagged samples and which is the 80% of the instances of the WQI-BP dataset. Evaluation of the prediction models is carried out to check the efficiency of the model using the metrics like R2

score, root mean squared error, mean squared error, and mean absolute error with the test data set.

Results of RNN Based WQI prediction model

The performance of the RNN-based WQI prediction model (RNN-WQI-BP model) is assessed by conducting experiments with a range of epochs, varying from 20 to 500. At 20 epochs, the model exhibits an MAE of 0.647, MSE of 0.662, RMSE of 0.8136, and R2 score of 0.584. With the progression to 50 epochs, a noticeable reduction is observed in MAE of 0.638, MSE of 0.642, RMSE of 0.8012, and a slight improvement in R2 score of 0.589. Continuing the training process to 100 epochs results in further refinement of predictions, as indicated by the diminishing MAE of 0.631, MSE of 0.628, and RMSE of 0.7925 values, accompanied by a modest increase in R2 score of 0.593.

As the training continues, particularly at 150 epochs, the predictive accuracy is enhanced, demonstrated by the decreasing MAE of 0.627, MSE of 0.612, and RMSE of 0.7823 values, with a corresponding improvement in R2 score of 0.606. The trend of refinement persists at 200 epochs, where the model attains lower MAE of 0.614, MSE of 0.584, and RMSE of 0.7642 values, contributing to a higher R2 score of 0.627. At the maximum epoch value of 500, the model predictive prowess becomes evident, as demonstrated by the notably reduced MAE of 0.598, MSE of 0.523, RMSE of 0.7232, and a favourable R2 score of 0.642. The performance evaluation of the WQI prediction model based on the WQI-BP dataset with traditional RNN-WQI-BP is shown in Table 2 and depicted in Fig.4.

Table 2. Prediction Results of RNN-WQI-BP Model for Various Epochs

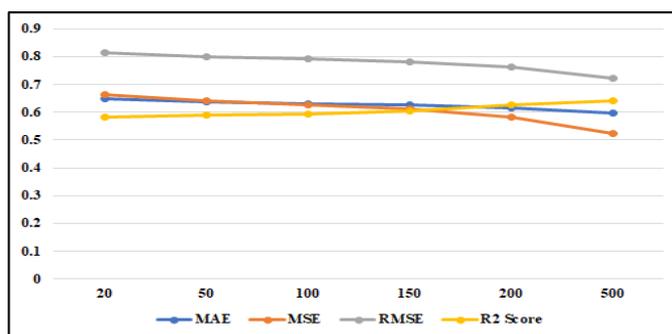| Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-BP	20	0.647	0.662	0.8136	0.584
	50	0.638	0.642	0.8012	0.589
	100	0.631	0.628	0.7925	0.593
	150	0.627	0.612	0.7823	0.606
	200	0.614	0.584	0.7642	0.627
	500	0.598	0.523	0.7232	0.642

Fig.4. Prediction Results of RNN-WQI-BP Model for Various Epochs

Results of LSTM Based WQI prediction model

The performance of the LSTM-based WQI prediction model (LSTM-WQI-BP model) is assessed by conducting experiments with a range of epochs, varying from 20 to 500. At 20 epochs, the model exhibits a MAE of 0.591, MSE of 0.613, RMSE of 0.7829, and an R2 score of 0.592. As the training process advances to 50 epochs, a gradual enhancement is observed in the model performance, leading to decreased MAE of 0.585, MSE of 0.596, and RMSE of 0.7720, accompanied by a marginally improved R2 score of 0.604. Continuing the training to 100 epochs demonstrates further refinement in predictive accuracy, highlighted by a lower MAE of 0.563, MSE of 0.583, and RMSE of 0.7635, while showcasing a notable improvement in the R2 score of

0.627. At 150 epochs, the model continues to progress, with reduced MAE of 0.552, MSE of 0.547, and RMSE of 0.7396, leading to an elevated R2 score of 0.636. As the training iterations reach 200 epochs, reflecting decreased MAE of 0.543, MSE of 0.535, and RMSE of 0.7314, resulting in a higher R2 score of 0.657. Finally, at 500 epochs, the model reaches a noteworthy level of predictive accuracy, evidenced by lower MAE of 0.524, MSE of 0.513, and RMSE of 0.7162, culminating in an impressive R2 score of 0.684. The results prove that increasing the number of epochs leads to better performance of the LSTM model in predicting the WQI-BP values. The performance evaluation of the LSTM model is shown in Table 3 and illustrated in Fig.5.

Table 3. Prediction Results of LSTM-WQI-BP Model for Various Epochs

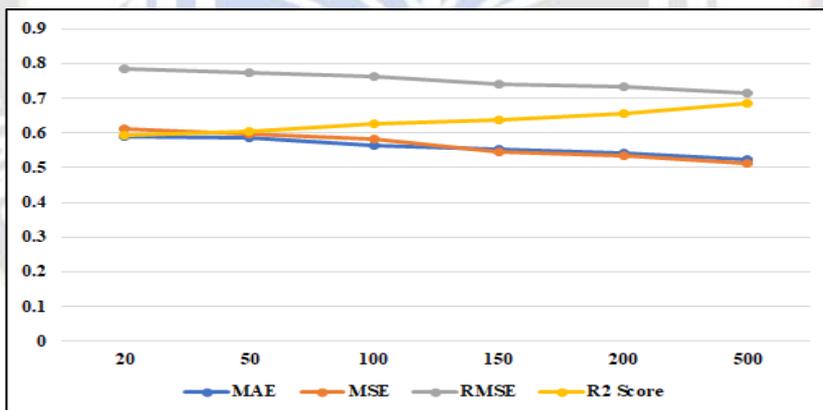| Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-BP	20	0.591	0.613	0.7829	0.592
	50	0.585	0.596	0.7720	0.604
	100	0.563	0.583	0.7635	0.627
	150	0.552	0.547	0.7396	0.636
	200	0.543	0.535	0.7314	0.657
	500	0.524	0.513	0.7162	0.684

Fig.5. Prediction Analysis of LSTM-WQI-BP Model for Various Epochs

Results of GRU Based WQI prediction model

The performance of the GRU-based WQI prediction model (GRU-WQI-BP model) is assessed by conducting experiments with a range of epochs, varying from 20 to 500. At 20 epochs, the model exhibits specific performance metrics, including a MAE of 0.612, MSE of 0.617, RMSE of 0.7855, and an R2 score of 0.591. As the training advances to 50 epochs, the model demonstrates incremental improvement, leading to lower MAE of 0.605, MSE of 0.59, and RMSE of 0.7681, accompanied by a higher R2 score of 0.618. Subsequently, with 100 epochs, the model's

performance continues to enhance, resulting in a reduced MAE of 0.572, while maintaining a similar MSE of 0.571 and exhibiting a lower RMSE of 0.7556 alongside an elevated R2 score of 0.624. The trend of progress is maintained at 150 epochs, where the model showcases further refinement in its predictions, leading to a decreased MAE of 0.564, lower MSE of 0.552, and RMSE of 0.7430, and aR2 score of 0.645. Continuing to 200 epochs, the model consistently improves its predictive accuracy, as reflected by a decreased MAE of 0.547, lower MSE of 0.531, and RMSE of 0.7287, coupled with an increased R2 score of 0.651.

Finally, at 500 epochs, the model attains a notable level of predictive excellence, with reduced MAE of 0.536 and MSE of 0.527, further showcasing a lower RMSE of 0.7259, culminating in an impressive R2 score of 0.672. The performance of the GRU-WQI-BP prediction with various epochs is shown in Table 4 and depicted in Fig.6.

Table 4. Prediction Results of GRU-WQI-BP Model for Various Epochs

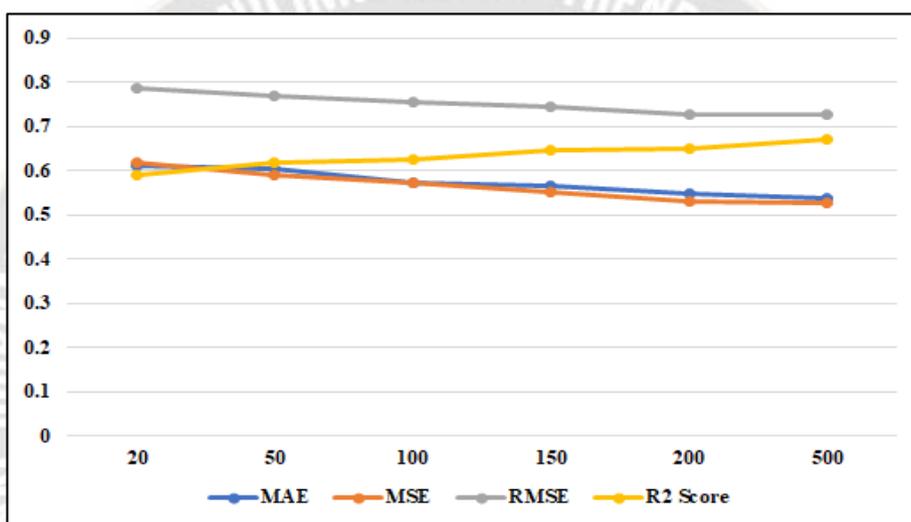| Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-BP	20	0.612	0.617	0.7855	0.591
	50	0.605	0.59	0.7681	0.618
	100	0.572	0.571	0.7556	0.624
	150	0.564	0.552	0.7430	0.645
	200	0.547	0.531	0.7287	0.651
	500	0.536	0.527	0.7259	0.672

Fig.6. Prediction Analysis of GRU-WQI-BP Model for Various Epochs

Results of TFT Based WQI prediction model

In the case of TFT based WQI model, the hyperparameter setting for TFT forecasting involved exploring a prediction time step of 30, setting the encoder layer to 4, using a fixed batch size of 64, varying state sizes from 32 to 256, and setting it to 64, trying out learning rates from 0.0001 to 0.1, varying the number of attention heads from 1 to 8, applying dropout rates from 0 to 0.4. The hyperparameter settings of the TFT-WQI-BP model is tabulated in Table 5.

Table 5: Setting of Special Hyperparameters for TFT Training

Time steps	Encoders layers	Batch sizes	State size	Learning rates	Attention heads	Dropout rate	Loss Function a	Loss Function b	Loss Function g
30	4	64	64	0.01	4	0.20, 0.30	0.80	0.01	0.10

The performance of the TFT-based WQI prediction model (TFT-WQI-BP model) is assessed by conducting experiments with a range of epochs, varying from 20 to 500. At 20 epochs, the model exhibits specific performance metrics, including a MAE of 0.532, MSE of 0.516, RMSE of 0.7183, and an R2 score of 0.623. As the training progresses to 50 epochs, the model showcases incremental refinement, manifesting in reduced MAE of 0.518, lower MSE of 0.498, and a decreased RMSE of 0.7057, accompanied by an elevated R2 score of 0.635. Subsequently, at 100 epochs, the model performance continues to advance, leading to a further reduced MAE of 0.507, with corresponding decreases in both MSE of 0.482 and RMSE of 0.6943, ultimately culminating in an enhanced R2 score of 0.646. As the model is further trained with 150 epochs, it continues to demonstrate improved predictive accuracy, reflected in the reduction of MAE of 0.481, lower MSE of 0.465, and a diminished RMSE of 0.6819, along with an elevated R2 score of 0.662.

833

Continuing the training process to 200 epochs, the model consistently enhances its predictive capability, resulting in a notable decrease in MAE of 0.448, MSE of 0.451, and RMSE of 0.6716, along with a substantial increase in R2 score of 0.687. Finally, with 500 epochs, the model achieves a commendable level of predictive excellence, as evidenced by the decreased MAE of 0.407 and MSE of 0.436, coupled with a further reduction in RMSE 0.6603, culminating in an impressive R2 score of 0.705. The performance of the TFT-WQI-BP prediction with various epochs is shown in Table 6 and illustrated in Fig.7.

Table 6. Prediction Results of TFT-WQI Model with WQI-BP Dataset for various Epochs

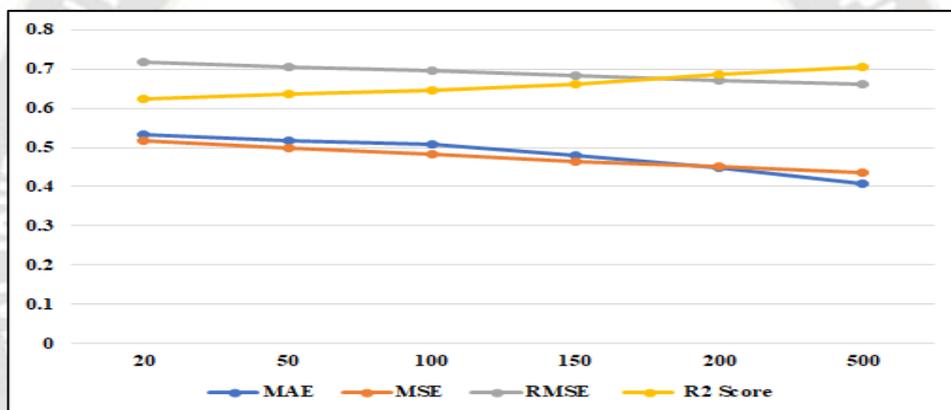| Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-BP	20	0.532	0.516	0.7183	0.623
	50	0.518	0.498	0.7057	0.635
	100	0.507	0.482	0.6943	0.646
	150	0.481	0.465	0.6819	0.662
	200	0.448	0.451	0.6716	0.687
	500	0.407	0.436	0.6603	0.705

Fig.7. Prediction Analysis of TFT-WQI Model with WQI-BP Dataset for Various Epochs

Comparative Analysis

The prediction results of WQI models for various epochs and dropouts have been observed while implementing deep learning algorithms to discover the best prediction results. It is proved that the models trained with 500 epochs with other hyperparameters such as adam optimizer, momentum as 0.8, dropout as 0.3 and activation function as relu for RNN, LSTM, GRU, and architecture TFT, produced the best results and are shown in Table 7 and depicted in Fig. 4.

Table 7. Overall Performance Results of WQI Models for Bharathapuzha Data

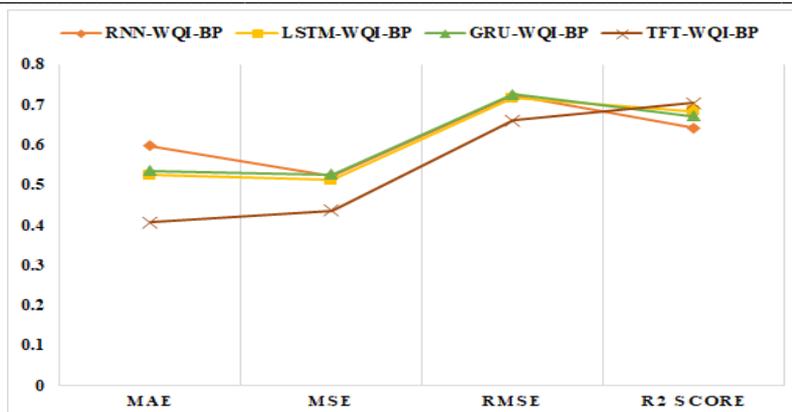| Dataset	Epoch	Models	MAE	MSE	RMSE	R2 Score
WQI-BP	500	RNN-WQI-BP	0.598	0.523	0.7232	0.642
		LSTM-WQI-BP	0.524	0.513	0.7162	0.684
		GRU-WQI-BP	0.536	0.527	0.7259	0.672
		TFT-WQI-BP	**0.407**	**0.436**	**0.6603**	**0.705**

Fig.8. Prediction Performance of Deep Learning Algorithms with Bharathapuzha Data

From the above results, it is observed that the TFT-based WQI prediction model shows promising results with a high R2 score value and less error rate. The mean absolute error for TFT based forecasting model is found less as compared to RNN, LSTM and GRU algorithms. The root mean squared error is observed to be less for the TFT-WQI-BP model when compared with RNN-WQI-BP, LSTM-WQI-BP and GRU-WQI-BP prediction model results. The R2 score value defines the accuracy of the model and is observed to be high for the TFT-WQI-BP forecasting model compared with other prediction models.

The investigations made in this research proved that the deep learning approach is useful for developing predictive models like water quality index prediction. It is confirmed that the recent deep learning approach improves the prediction accuracy of different WQI predictive models. Through feature selection, the association between the pool of predictors and the targeted variable is strengthened which enables deep neural network architectures TFT, GRU, LSTM, and RNN to improve the learning of trends in the data. The prediction rate of WQI models is increased through learning the self-extracted features in TFT, GRU, LSTM, and RNN networks. The error rate of trained models is decreased by properly configuring the hyperparameters during network training.

5. CONCLUSION

This study demonstrated the importance of temporal fusion transformer architecture with the Bharathapuzha dataset containing physiochemical and seasonal attributes. The application of deep learning architectures for river water quality time series forecasting was attempted to prove that deep learning is an effective approach for accurate WQI prediction. The data was collected from the sampling stations and visual crossing site during the period 2020 to 2021 and a time series dataset was developed. The river water quality forecasting model has been designed and developed using deep learning architectures such as RNN, LSTM, GRU and

TFT. The performance of the TFT-WQI-BP model was evaluated and compared with the prediction results of models trained with other deep learning models. From the evaluation results, it is observed that the TFT-based WQI prediction model provides an enhanced and efficient WQI prediction model. The developed model can even be used as pre-trained models with transfer learning to improve the efficiency of the prediction model.

REFERENCES

[1] Gopal Krishan, Surjeet Singh, Kumar CP, Suman Gurjar and Ghosh NC, (2016), "Assessment of Water Quality Index (WQI) of Groundwater in Rajkot District, Gujarat, India", Journal of Earth Science & Climatic Change, Volume 7, Issue 3, 1000341 (2016)

[2] Umair Ahmed, Rafia Mumtaz, Hirra Anwar, Asad A. Shah, Rabia Irfan and José García-Nieto (2019), 'Efficient Water Quality Prediction Using Supervised Machine Learning'- Journal of Water.

[3] M. Ezhilarasi and V. Senthilkumar (2018), Geo-Chemical Analysis for Groundwater Quality Using Geospatial Application, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04.

[4] Shuangyin Liu, Haijiang Taia, Qisheng Dinga, Daoliang Lia, Longqin Xub, Yaoguang Weia (2011), 'A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction'- Journal of Elsevier.

[5] Salisu Yusuf Muhammad, Mokhairi Makhtar, Azilawati Rozaimee, Azwa Abdul Aziz and Azrul Amri Jamal (2015), 'Classification Model for Water Quality using Machine Learning Techniques' - International Journal of Software Engineering and Its Applications Vol. 9, No. 6, pp. 45-52.

[6] Li L, Jiang P, Xu H, Lin G, Guo D, Wu H (2019) Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang

River, China. Environ Sci Pollut Res 26(19): 19879–19896

[7] Liu J, Yu C, Hu Z, Zhao Y, Bai Y, Xie M, Luo J (2020) Accurate prediction scheme of water quality in smart mariculture with deep Bi-S-SRU learning network. IEEE Access 8:24784–24798

[8] Yahya A, Saeed A, Ahmed AN, Binti Othman F, Ibrahim RK, Afan HA, Elshafie A (2019) Water quality prediction model-based support vector machine model for ungauged river catchment under dual scenarios. Water 11(6):1231

[9] J. P. Nair and M. S. Vijaya, "Predictive Models for River Water Quality using Machine Learning and Big Data Techniques - A Survey," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 1747-1753.

[10] Nair, Jitha P., and M. S. Vijaya. 'River Water Quality Prediction and Index Classification Using Machine Learning'. *Journal of Physics: Conference Series*, vol. 2325, no. 1, Aug. 2022, p. 012011.

[11] Heddam, S., 2014. Generalized regression neural network-based approach for modelling hourly dissolved oxygen concentration in the Upper Klamath River, Oregon, USA. Environmental Technology (United Kingdom) 35, 1650–1657. 10.1080/09593330.2013.878396.

[12] Nair, J.P., Vijaya, M.S. (2023). Exploratory Data Analysis of Bhavani River Water Quality Index Data. In: Kumar, S., Hiranwal, S., Purohit, S.D., Prasad, M. (eds) Proceedings of International Conference on Communication and Computational Technologies. Algorithms for Intelligent Systems. Springer, Singapore.

[13] Santhana Lakshmi, V., Vijaya, M.S. (2022). A Study on Machine Learning-Based Approaches for PM2.5 Prediction. In: Karrupusamy, P., Balas, V.E., Shi, Y. (eds) Sustainable Communication Networks and Application. Lecture Notes on Data Engineering and Communications Technologies, vol 93. Springer, Singapore.

[14] Jitha P Nair, and Vijaya M S. 'Design and Development of Efficient Water Quality Prediction Models Using Variants of Recurrent Neural Networks'. European Chemical Bulletin, vol. 12, no. Special Issue 5, May 2023.

[15] Heddam, S., Kisi, O., 2018. Modelling daily dissolved oxygen concentration using least square support vector machine, multivariate adaptive regression splines and M5 model tree. J. Hydrol. 559, 499–509.

[16] Nair, J. P ., & M. S., V. . (2023). Temporal Fusion Transformer: A Deep Learning Approach for Modeling and Forecasting River Water Quality Index. International Journal of Intelligent Systems and Applications in Engineering, 11(10s), 277–293.