PSGR Krishnammal College for Women

TEMPORAL FUSION TRANSFORMER AND TRANSFER LEARNING FOR TIME SERIES RIVER WATER QUALITY PREDICTION MODEL

Thesis submitted to

BHARATHIAR UNIVERSITY

for the award of the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER SCIENCE

By

JITHA P NAIR

Under the guidance of

Dr. (Mrs.) M.S. VIJAYA M.Sc., M.Phil., Ph.D.

Associate Professor and Head (Retd.)

Department of Computer Science (PG)

PSGR Krishnammal College for Women

Peelamedu, Coimbatore - 641 004, India

www.psgrkcw.ac.in

SEPTEMBER 2023

8. HETEROGENEOUS TRANSFER LEARNING FOR WQI PREDICTION MODELS

The implementation of homogenous transfer learning approach to build WQI model was considered in the previous chapter as the samples of Bharathapuzha river and the Bhavani River contain the same set of water quality parameters and the datasets used in pre-trained WQI model and the TL based WQI modes have similar attribute set. It is observed from the literature that two parameters, flow rate and Sodium Absorption Ratio (SAR) are considered to be important in determining water quality index. But these two parameters are observed in the monitoring stations of Bhavani River. SAR and flow rate have a significant impact on the calculation of the WQI. The sodium Absorption Ratio measures the relative concentrations of sodium, calcium, and magnesium ions in water and provides insights into the risk of soil degradation, reduced water infiltration, and overall soil health. Flow rate refers to the volume of water passing through a specific point within a given time period. An optimal flow rate ensures sufficient exposure for physical, chemical, and biological reactions, promoting the effective removal of contaminants and pollutants. Hence an attempt is made to extend the time series data of the Bharathapuzha river by incorporating these two parameters and to apply the heterogeneous transfer learning for the development of a more robust WQI prediction model.

Heterogeneous transfer learning refers to a scenario where the feature spaces of the source and target domains exhibit notable differences. This variant of transfer learning presents greater complexity compared to homogeneous transfer learning, as it necessitates the establishment of connections between these distinct feature spaces. Now the extended Bharathapuzha river dataset has different feature set compared with Bhavani River dataset.

This chapter delves into the conceptualization and implementation of a heterogeneous transfer learning approach for building WQI prediction models, shedding light on the intricate interplay between distinct datasets and their impact on model performance. Through a comprehensive analysis of heterogeneous transfer learning techniques, this chapter elucidates how harnessing knowledge from disparate sources can lead to a more robust and reliable prediction model. The methodology of the heterogeneous transfer learning approach applied to build WQI prediction model using extended Bharathapuzha River dataset and the TFT based pre-trained model is elucidated in this chapter.

WQI PREDICTION MODEL USING HETEROGENEOUS TRANSFER LEARNING WITH TFT PRE-TRAINED MODEL

The aim of this work is to enhance the efficacy of the WQI model using heterogenous transfer learning (HTL) and training the extended Bharathapuzha River data. The enhancement is achieved through the implementation heterogenous transfer learning approach with TFT based pre-trained model is used. This work comprehensively employs a spectrum of deep learning architectures including RNN, LSTM, GRU, and the TFT architecture. The reason for using RNN, LSTM, and GRU simultaneously when the TFT architecture is also employed in heterogeneous transfer learning-based training is to study and confirm which deep neural network performs better when a heterogenous transfer learning approach is adopted. This empirical exploration gains depth through the incorporation of HTL, enhancing the holistic and inventive framework for building WQI prediction model.

Methodology

Leveraging the advantages of transfer learning approaches, the integration of pre-existing knowledge from extensively trained models on vast datasets empowers effective performance, even when confronted with limited training data for particular tasks. In this context, due to the lack of an existing pre-trained model for the implementation of the transfer learning approach, the developed model approach is employed. The proposed heterogenous transfer learning-based WQI prediction model is constructed using three essential components such as (i) Extended Bhrathapuzha River Dataset (ii) TFT based pre-trained WQI model (iii) Networks for heterogeneous transfer learning. The heterogeneous transfer learning-based WQI prediction models are built using RNN, LSTM, GRU and also using TFT, using the WQI-EBP dataset and the TFT based pre-trained model as depicted in Fig. 8.1.

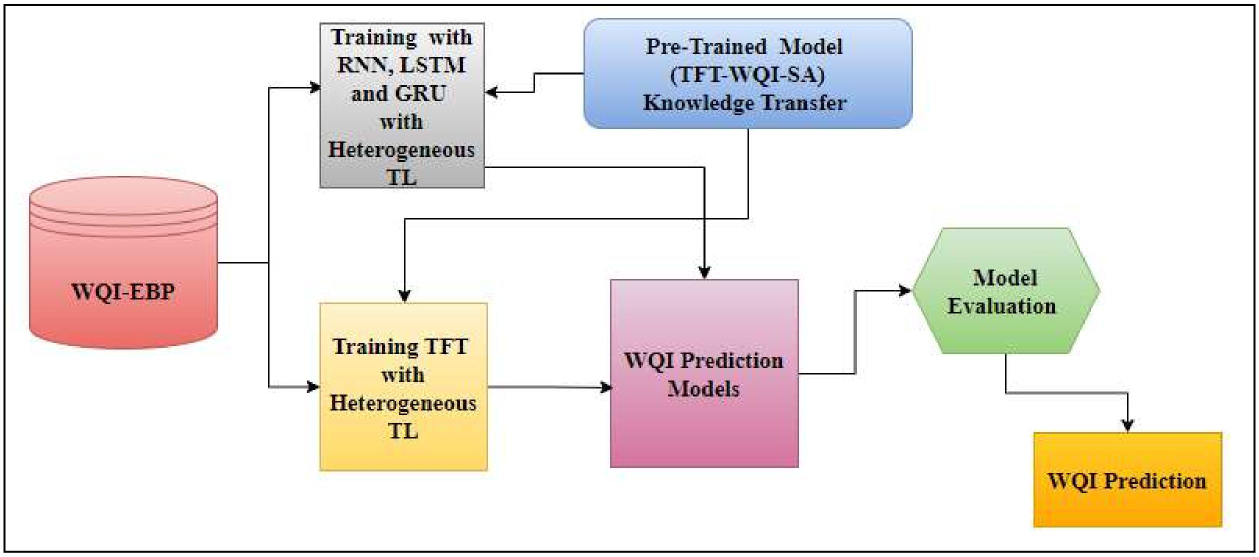

Fig. 8.1. The Methodology of Heterogenous Transfer Learning based TFT Pre-trained Model

Extended Bhrathapuzha River Dataset (WQI-EBP)

The river water quality samples comprising 26 physicochemical parameters, 10 seasonal parameter, 3 spatial attributes, temporal attribute and 2 calculated parameters such as SAR and flow rate of Bharathapuzha river observed for the period January 2020 to December 2021 are taken.

Flow rate refers to the volume of water passing through a specific point within a given time period. SAR measures the relative concentrations of sodium, calcium, and magnesium ions in water and provides insights into the risk of soil degradation, reduced water infiltration. The spatial parameters such as longitude, latitude, station ID, and temporal parameter date, are added to the time series data. The water quality index value for each sample is calculated resulting in time series data with a total of 43 attributes and 2190 instances.

Exploratory Data Analysis is done on the raw collected data to gain insights into its characteristics and assess the significance of each parameter in determining the WQI. Various statistical techniques, including heatmap analysis, boxplot analysis, pair plot analysis, and histogram analysis, are used to study the distribution of parameter values. The min-max normalization is implemented on water quality parameters, to standardize the parameter values. Through the select K best feature selection method, the river water quality dataset is improved, resulting in the extended Bharathapuzha river dataset with 2190 instances and 41 features along with calculated WQI, which is named as WQI-EBP dataset as indicated in Table XII of Chapter 3.

Pre-trained Model

The WQI models developed using Bhavani River data i.e., the WQI-SA dataset and trained with the TFT, as described in Chapter 6 are considered here as a pre-trained model. The optimal performance of TFT based WQI model has been achieved when trained with 500 epochs and a dropout rate of 0.3. The TFT-WQI-SA model demonstrated superior R2 scores and lower error rates as per the results given in Table. XXXII of Chapter 6, which is reproduced in Table LVII.

Table LVII. Prediction Results of TFT based Pre-trained Model Based on WQI-SA Dataset

Model	Dropout	Epoch	MAE	MSE	RMSE	R2 Score
TFT-WQI-SA	**0.3**	**500**	**0.122**	**0.167**	**0.4087**	**0.941**
RNN -WQI-SA			0.428	0.384	0.6197	0.82
LSTM-WQI-SA			0.298	0.2084	0.4565	0.856
GRU-WQI-SA			0.39	0.2149	0.4636	0.839

Hence the TFT-WQI-SA model, developed for the prediction of water quality index, is used as the pre-trained model in this work to facilitate heterogeneous transfer learning. The weights of the parameters can be transferred while training Bharathapuzha River water data using transfer learning to achieve a better prediction model.

Model Building using Heterogenous Transfer Learning

Here, to build heterogeneous transfer learning based WQI prediction models, the architectures RNN, LSTM, GRU and TFT are used independently to learn the water quality patterns from WQI-EBP dataset. RNN, LSTM, and GRU have been established as architectures for sequence data training, recognized for their impressive precision. TFT stands out in capturing temporal relationships within time series data, enabling adept recognition of temporal patterns and trends through transformer-based attention mechanisms.

Heterogeneous transfer learning involves a multi-step process that enhances the prediction capabilities of a model by transferring knowledge from one domain to improve predictions in another domain. In this context, the process begins with the utilization of a pre-existing model, TFT-WQI-SA, as the initial foundation. In this starting phase, all layers of the model, except the output layer, are maintained in a frozen state. This ensures the retention of the insights and understanding accumulated during the training of the original model on the Bhavani River dataset. By preserving this acquired knowledge, the model gains a preliminary understanding of water quality dynamics.

Subsequently, the process advances to the retraining phase, where the frozen model is fine-tuned using the WQI-EBP dataset, which pertains to the Bharathapuzha River. During this step, the model parameters are adjusted to align with the unique attributes and characteristics of the Bharathapuzha River water quality. This fine-tuning process allows the model to adapt its predictive capabilities to the specific conditions of the target domain. The pre-trained knowledge inherited from the TFT-WQI-SA model serves as a foundational framework that guides learning process of the model during this phase.

Effective execution of heterogeneous transfer learning relies on strategic hyperparameter adjustments, which yield significant influence over the process. Precise tuning of hyperparameters, including the learning rate, batch size, epochs, activation functions, and regularization strength, holds the key to enhancing the performance of the model. Alterations to the learning rate dictate the convergence speed of the model, ensuring a proficient and impactful learning trajectory. Manipulating batch sizes impacts weight updates, potentially aiding the model in avoiding local optima and bolstering its ability to generalize insights. Modulating the number of epochs governs training depth, striking a harmonious equilibrium between underfitting and overfitting. The selection of activation functions shapes the adeptness of the model at capturing intricate patterns and gradients.

In case of training the TFT network for heterogeneous transfer learning, the special hyperparameters such as time step, the encoder layer, batch sizes, state sizes, learning rates, number of attention heads, dropout rates, and Loss Function values are configured appropriately while training the TFT. This meticulous parameter tuning within the TFT training process is paramount to achieve effective and adaptable knowledge transfer while learning the patterns.

Thus, four independent models are built using RNN, LSTM, GRU and the architecture TFT. The models are named as RNN-WQI-EBP-HTL, LSTM-WQI-EBP-HTL, GRU-WQI-EBP-HTL and TFT-WQI-EBP-HTL. The performance of predictive models is assessed using evaluation metrics including MAE, MSE, RMSE, and R2 scores, with 20% of instances of the WQI-EBP dataset as test set.

Experiments and Results

The empirical investigations are conducted through training the extended Bharathapuzha River water dataset alongside the pre-trained TFT-WQI-SA model with TFT and RNN variants, and are executed using Python libraries. The training dataset encompassed 1752 labelled instances extracted from the WQI-EBP dataset. Each of the distinct architectures, namely RNN, LSTM, GRU, and the TFT architecture, is independently implemented. The evaluation of the predictive models as an assessment of their efficacy through metrics such as MAE, MSE, RMSE, and R2 scores, employing a separate test dataset comprising 438 instances.

The RNN and heterogenous transfer learning based WQI prediction model (RNN-WQI-EBP-HTL model) trained using the pretrained TFT-WQI-SA model and WQI-EBP dataset have been experimented with various epochs such as from 20 to 500 where various metrics are measured at different epochs. The hyper parameter setting for RNN, LSTM and GRU is done as mentioned in the previous section with Table XLI.

Initiating with 20 epochs, the model yields an MAE of 0.4, which reflects the average magnitude of prediction errors. Simultaneously, the MSE and RMSE values are registered at 0.401 and 0.6332, respectively, and the R2 Score, indicative goodness fit of the model, attains a value of 0.83. Progressing to 50 epochs, the model exhibits a similar MAE of 0.398, accompanied by a slightly reduced MSE and RMSE, measuring 0.391 and 0.6253, respectively. The R2 Score continues to show a favourable value of 0.836. As the epoch count increases to 100, there is an improvement in the performance of the model, as indicated by the diminished MAE of 0.327. Both the MSE and RMSE values experience a reduction to 0.357 and 0.5975, respectively, while the R2 score maintains an elevated level of 0.845, demonstrating the robust predictive capabilities of the model. Advancing to 150 epochs, the model further refines its performance, yielding an improved MAE of 0.282. The MSE and RMSE continue their decreasing trend, reaching values of 0.346 and 0.5882, respectively and the R2 Score rises to 0.856.

At 200 epochs, the model continues to enhance its predictive accuracy, with the MAE dropping to 0.224. The MSE and RMSE follow suit, settling at 0.264 and 0.5138, respectively. The R2 Score remains high at 0.862, affirming the strong fit of the model to the dataset. Lastly, with 500 epochs, the model attains its optimal performance, as evidenced by an MAE of 0.219, the lowest among the examined epochs. The MSE and RMSE reach their respective minima of 0.159 and 0.3987, highlighting the precision in prediction. The R2 Score reaches an impressive value of 0.88, indicating the exceptional model fit to the data. The performance of RNN-WQI-EBP-HTL prediction model is depicted in Table LVIII.

Table LVIII. Performance of RNN-WQI-EBP-HTL Prediction Model for Different Epochs

Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-EBP	20	0.4	0.401	0.6332	0.83
	50	0.398	0.391	0.6253	0.836
	100	0.327	0.357	0.5975	0.845
	150	0.282	0.346	0.5882	0.856
	200	0.224	0.264	0.5138	0.862
	500	0.219	0.159	0.3987	0.88

The results of the LSTM and heterogenous transfer learning based WQI prediction model (LSTM-WQI-EBP-HTL model) trained using the pretrained TFT-WQI-SA model and WQI-EBP dataset have been experimented with various epochs such as from 20 to 500 where various metrics are measured at different epochs. Commencing with 20 epochs, the model achieves an MAE of 0.234, characterizing the average magnitude of prediction errors. Concurrently, the MSE and RMSE values stand at 0.226 and 0.4754, respectively, and the R2 score attains a commendable value of 0.89, indicative of its capacity to explain data variance. Advancing to 50 epochs, the model maintains its performance with a similar MAE of 0.225, accompanied by a slightly reduced MSE and RMSE, which are recorded at 0.215 and 0.4637, respectively and the R2 Score ascends to 0.895. Upon reaching 100 epochs, the model refines its performance further, achieving an MAE of 0.213. The MSE and RMSE values continue to diminish, settling at 0.21 and 0.4583, respectively. The R2 Score maintains an elevated level of 0.91, signifying the strength of the model.

Advocating for 150 epochs, the model continues to enhance its performance, as evidenced by a reduced MAE of 0.19. Both the MSE and RMSE values further decrease, reaching values of 0.194 and 0.4405, respectively and the R2 Score achieves to 0.916. Advancing to 200 epochs, the model continues its trajectory of heightened performance, achieving an even lower MAE of 0.167. The MSE and RMSE values exhibit a parallel decline, settling at 0.182 and 0.4266, respectively. The R2 Score maintains an impressive level of 0.92. Finally, at 500 epochs, the model attains its pinnacle performance, characterized by a minimal MAE of 0.138, reflecting heightened prediction accuracy. The MSE and RMSE values follow suit, reaching their nadir at 0.168 and 0.4099, respectively and the R2 score attains an elevated value of 0.942. The performance of LSTM-WQI-EBP-HTL prediction model is depicted in Table LIX.

Table LIX. Performance of LSTM-WQI-EBP-HTL Prediction Model for Different Epochs

Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-EBP	20	0.234	0.226	0.4754	0.89
	50	0.225	0.215	0.4637	0.895
	100	0.213	0.21	0.4583	0.91
	150	0.19	0.194	0.4405	0.916
	200	0.167	0.182	0.4266	0.92
	500	0.138	0.168	0.4099	0.942

The results of the GRU and heterogenous transfer learning based WQI prediction model (GRU-WQI-EBP-HTL model) trained using the pretrained model TFT-WQI-SA model and WQI-EBP dataset have been experimented with various epochs such as from 20 to 500 where various metrics are measured at different epochs. Commencing with 20 epochs, the model exhibits an MAE of 0.278, signifying the average magnitude of prediction errors. Simultaneously, the MSE and RMSE values stand at 0.261 and 0.5109, respectively, and the R2 Score attains a commendable value of 0.872. Advancing to 50 epochs, the model sustains its performance, attaining an MAE of 0.262, accompanied by a slightly diminished MSE and RMSE, quantified at 0.243 and 0.4930, respectively. Notably, the R2 Score further elevates to 0.891. As the epoch count increases to 100, the predictive accuracy of the model refines, yielding an MAE of 0.255. The MSE and RMSE values continue their downward trend, registering at 0.218 and 0.4669, respectively. The R2 Score rises to 0.902.

Upon reaching 150 epochs, the model continues its trajectory of improvement, as evidenced by a reduced MAE of 0.216. Both the MSE and RMSE values further decrease, settling at 0.197 and 0.4438, respectively. The R2 Score advances to 0.91, signifying the robustness of the model. Advocating for 200 epochs, the model continues to enhance its predictive capabilities, achieving a diminished MAE of 0.193. The MSE and RMSE values exhibit parallel reduction, reaching values of 0.189 and 0.4347, respectively. The R2 Score maintains an elevated level of 0.921. Finally, at 500 epochs, the model attains its peak performance, characterized by a minimal MAE of 0.185, signifying heightened prediction accuracy. The MSE and RMSE values follow suit, reaching their nadir at 0.175 and 0.4183, respectively. Impressively, the R2 Score attains an elevated value of 0.935, the emblematic remarkable ability of the model to capture and explain data variance. The performance of GRU-WQI-EBP-HTL prediction model is depicted in Table LX.

Table LX. Performance of GRU-WQI-EBP-HTL Prediction Model for Different Epochs

Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-EBP	20	0.278	0.261	0.5109	0.872
	50	0.262	0.243	0.4930	0.891
	100	0.255	0.218	0.4669	0.902
	150	0.216	0.197	0.4438	0.91
	200	0.193	0.189	0.4347	0.921
	500	0.185	0.175	0.4183	0.935

In the case of TFT architecture implementation, the model employs the Adam, optimization technique and divides the dataset into distinct sets for training, validation, and testing.

This division allows for learning, hyperparameter tuning, and performance evaluation, respectively. The optimal model selection during training is facilitated through the utilization of early stopping. The exploration for improved TFT parameters encompasses a prediction time step of 30 steps, the encoder layer is set as 4, batch sizes are fixed to 64, state sizes from 32 to 256 and it is set to 64, learning rates from 0.0001 to 0.1, number of attention heads from 1 to 8, dropout rates from 0 to 0.4, Loss Function values from 0.1 to 0.9, Loss Function b values from 0.0001 to 0.5, and Loss Function g values from 0.1 to 0.5. The final parameters for the enhanced TFT of the two units are presented in Table LXI.

Table LXI. Special Hyperparameter for TFT

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	0.2, 0.3	0.80	0.01	0.10

The TFT and heterogenous transfer learning-based WQI prediction model (TFT-WQI-EBP-HTL model) trained using the pre-trained TFT-WQI-SA model and WQI-EBP dataset have experimented with various epochs such as from 20 to 500 where various metrics are measured at different epochs. Commencing with 20 epochs, the model exhibits an MAE of 0.257, denoting the average magnitude of prediction errors. Correspondingly, the MSE and RMSE values are recorded at 0.281 and 0.5301, respectively, representing the model's ability to minimize error and its prediction accuracy. The R2 Score, a measure of the model's fit to the data, stands at 0.89, indicating a favourable fit. As epochs are increased to 50, there is a slight improvement in the MAE to 0.215, accompanied by a decrease in MSE and RMSE to 0.269 and 0.5187, respectively. The R2 Score remains high at 0.91, affirming the model's capacity to explain variance in the data. Continuing to 100 epochs, the model's performance improves further, with the MAE reducing to 0.197. The MSE and RMSE also decrease to 0.252 and 0.5020, respectively, while the R2 Score continues to rise, reaching 0.923, indicating a strong correlation between predicted and actual values.

At 150 epochs, the model experiences MAE as a notable reduction to 0.142, while the MSE and RMSE continue to decrease, attaining values of 0.227 and 0.4764, respectively. The R2 Score remains consistently high at 0.932, showcasing the predictive prowess of the model. With 200 epochs, the MAE sees a marginal drop to 0.138, reflecting the refined predictive accuracy of the model. This is mirrored in the MSE and RMSE, which decrease to 0.183 and 0.4278, respectively. The R2 Score rises notably to 0.958, underscoring the strong fit to the dataset. Finally, at 500 epochs, the performance is at its peak, with the MAE and MSE reaching their lowest values of 0.132 and 0.138, respectively. The RMSE remains relatively low at 0.3715, and the R2 Score reaches its highest point of 0.974, signifying an excellent fit between predicted and actual values.

The results indicate that the TFT-WQI-EBP-HTL prediction model performs well in predicting the target variable. Higher epochs tend to yield better performance in terms of accuracy and precision, as reflected in lower MAE, MSE, and RMSE values, higher R2 Score, and lower quantile loss as depicted in Table LXII.

Table LXII. Performance of TFT-WQI-EBP-HTL Prediction Model for Different Epochs

Dataset	Epochs	MAE	MSE	RMSE	R2 Score
WQI-EBP	20	0.257	0.281	0.5301	0.89
	50	0.215	0.269	0.5187	0.91
	100	0.197	0.252	0.5020	0.923
	150	0.142	0.227	0.4764	0.932
	200	0.138	0.183	0.4278	0.958
	500	0.132	0.138	0.3715	0.974

Different dropout rates, specifically 0.2 and 0.3, are employed in a series of experiments aimed at constructing water quality index prediction models using the WQI-EBP dataset and the pre-trained from the TFT-WQI-SA. The outcomes of these experiments, for various evaluation metrics, are detailed in Table LXIII.

Table LXIII. Performance Results of Heterogenous TL Based WQI Prediction Models for Different Dropout Rates

Dataset	Models	Drop out	MAE	MSE	RMSE	R2 Score
WQI-EBP	RNN-WQI-EBP-HTL	0.3	0.219	0.159	0.3987	0.88
		0.2	0.282	0.346	0.5882	0.856
	LSTM-WQI-EBP-HTL	0.3	0.138	0.168	0.4099	0.942
		0.2	0.167	0.182	0.4266	0.92
	GRU-WQI-EBP-HTL	0.3	0.185	0.175	0.4183	0.935
		0.2	0.216	0.197	0.4438	0.91
	TFT-WQI-EBP-HTL	0.3	0.132	0.138	0.3715	0.974
		0.2	0.142	0.227	0.4764	0.932

The prediction results of HTL based WQI models developed by using the pre-trained model and the WQI-EBP dataset for various epochs and dropouts have been observed while implementing RNN, LSTM, GRU and the architecture TFT to discover the best prediction results. It is proved that the models trained with 500 epochs and dropout rate 0.3 with other hyperparameters like adam optimizer, momentum as 0.8, and activation function as relu produced the best results and are shown in Table LXIV and depicted in Fig. 8.2.

Table LXIV. Final Results of Heterogenous TL Based WQI Prediction Models

Pretrained Model	Models	MAE	MSE	RMSE	R2 Score
TFT-WQI-SA	**TFT-WQI-EBP-HTL**	**0.132**	**0.138**	**0.3715**	**0.974**
	RNN-WQI-EBP-HTL	0.219	0.159	0.3987	0.88
	LSTM-WQI-EBP-HTL	0.138	0.168	0.4099	0.942
	GRU-WQI-EBP-HTL	0.185	0.175	0.4183	0.935

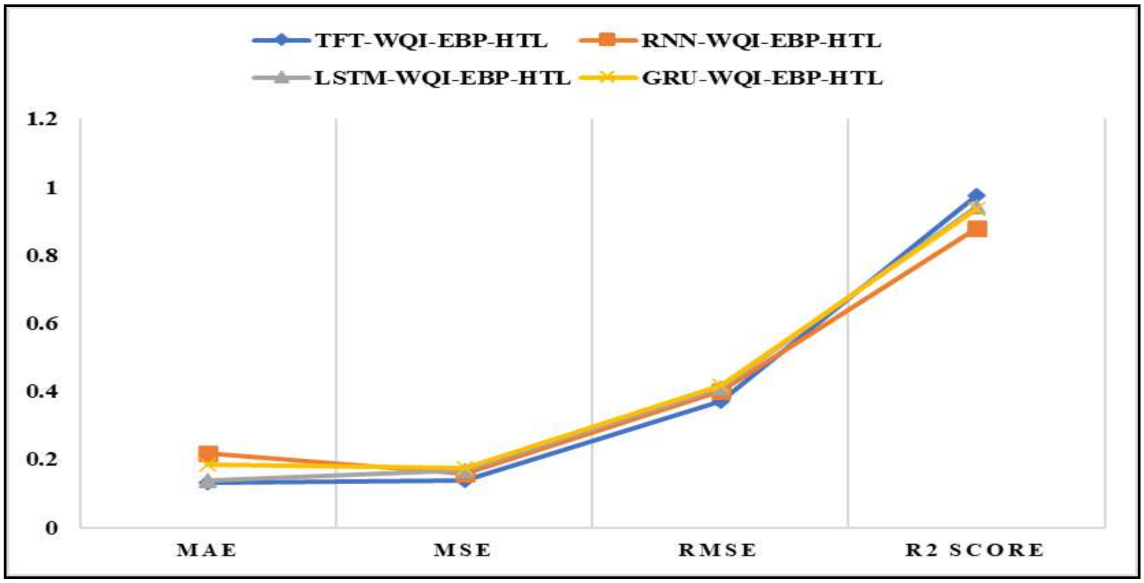

Fig.8.2. Prediction Performance of WQI prediction Models using Heterogenous TL

Comparative Analysis of Homogeneous Vs Heterogeneous Transfer Learning based WQI Prediction Models

The performance results of the WQI prediction models RNN-WQI-EBP-HTL, LSTM-WQI-EBP-HTL, GRU-WQI-EBP-HTL and TFT-WQI-EBP-HTL built using the pre-trained model (TFT-WQI-SA) and the extended WQI-EBP dataset using heterogeneous learning, are compared with homogenous transfer learning based WQI models RNN-WQI-BP-TFL, LSTM-WQI-BP-TFL, GRU-WQI-BP-TFL and TFT-WQI-BP-TFL described in section 7.3.

The comparative analysis of the WQI prediction models reveals distinct trends in performance across key evaluation metrics. Among the heterogeneous transfer learning models, TFT-WQI-EBP-HTL emerges as a standout performer, boasting the lowest MAE of 0.132, MSE of 0.138, RMSE of 0.3715 and R2 Score of 0.974 values. In contrast, the RNN-WQI-EBP-HTL model exhibits moderately elevated MAE of 0.219 and MSE of 0.159 values, accompanied by a higher RMSE of 0.3987 and lower R2 Score of 0.88 compared to the TFT-based counterpart. Similarly, LSTM-WQI-EBP-HTL demonstrates competitive performance, with a relatively low MAE of 0.138 and MSE 0.168, resulting in a favourable RMSE of 0.4099 and R2 Score of 0.942. GRU-WQI-EBP-HTL showcases a comparable trend, with an MAE of 0.185, MSE of 0.175, RMSE of 0.4183, and an impressive R2 Score of 0.935.

The TFT-WQI-BP-TFL model yields consistent results comparable to TFT-WQI-EBP-HTL, indicating the effectiveness of TFT-based architectures. In contrast, the RNN-WQI-BP-TFL model displays slightly lower accuracy with an MAE of 0.235, an MSE of 0.246, an RMSE of 0.4960, and an R2 Score of 0.86. The LSTM-WQI-BP-TFL model showcases a balance between accuracy and efficiency, with an R2 Score of 0.902. Finally, the GRU-WQI-BP-TFL model provides comparable results with an R2 Score of 0.892.

The comparative performance results of prediction models built with homogenous and heterogenous transfer learning are tabulated in Table LXIII and illustrated in Fig.8.3. Through this implementation of hetrogenous transfer learning, a holistic approach has been developed for building WQI prediction model and from the prediction results it is proved that the prediction models built with heterogeneous transfer learning yields better results. The TFT-WQI-EBP-HTL model demonstrated superior performance results in predicting the water quality index, outperforming the other models in terms of evaluation metrics MAE, MSE, RMSE, and R2 score.

Table LXV. Comparative Results of Heterogeneous and Homogenous TL Based WQI Prediction Models

Dataset	Pre-trained Model	Models	MAE	MSE	RMSE	R2 Score
WQI-EBP	TFT-WQI-SA	**TFT-WQI-EBP-HTL**	**0.132**	**0.138**	**0.3715**	**0.974**
		RNN-WQI-EBP-HTL	0.219	0.159	0.3987	0.88
		LSTM-WQI-EBP-HTL	0.138	0.168	0.4099	0.942
		GRU-WQI-EBP-HTL	0.185	0.175	0.4183	0.935
WQI-BP	TFT-WQI-SA	TFT-WQI-BP-TFL	0.132	0.138	0.3715	0.962
		RNN-WQI-BP-TFL	0.235	0.246	0.4960	0.86
		LSTM-WQI-BP-TFL	0.195	0.137	0.3701	0.902
		GRU-WQI-BP-TFL	0.215	0.137	0.3701	0.892

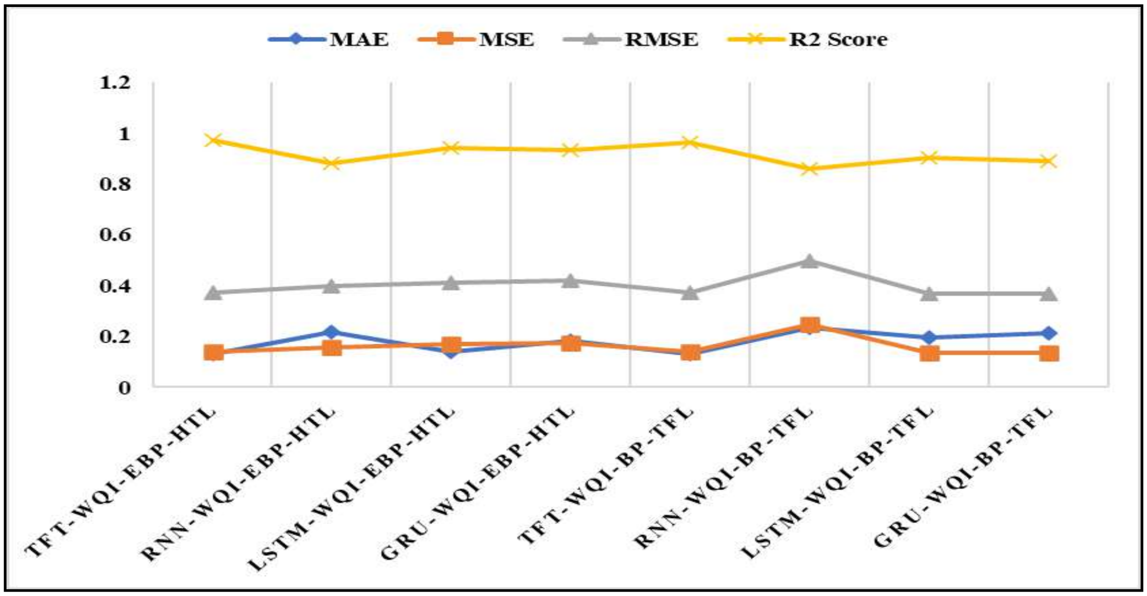

Fig.8.3. Performance Comparison Results of Heterogeneous and Homogenous TL based WQI Prediction Models

Findings

This work reveals that the integration of heterogeneous transfer learning significantly enhances the accuracy and efficiency of water quality prediction models. By leveraging pre-trained models and knowledge from one river dataset to enhance predictions for another river dataset, the approach demonstrates notable improvements across various evaluation metrics. By incorporating flow rate and SAR alongside existing water quality parameters, the dataset is expanded to provide a more comprehensive and understanding of water quality which enables a heterogeneous transfer learning approach to produce a more generalized WQI model. The highlight of this work is the effectiveness of temporal fusion transformers with hetrogenous transfer learning to predict WQI prediction. The prediction model benefits from the pre-trained knowledge, which helps to capture important features and patterns relevant to water quality. This transfer of knowledge enables the model to generalize better and make more accurate predictions. The integration of the Temporal Fusion Transformer further improves the prediction performance. The findings emphasize the potential of heterogeneous transfer learning as a powerful tool for improving water quality prediction models, particularly evident in the superior performance of the TFT-based approach. This highlights the robustness of the predictive model, its capabilities and its generalization across distinct water bodies.

SUMMARY

The chapter commenced by elucidating the significance of water quality prediction and the constraints posed by limited training data. Through the integration of flow rate and SAR in conjunction with existing water quality parameters, the Bharathapuzha river dataset is extended and their importance is highlighted. The chapter highlighted the need for innovative approaches to mitigate this limitation, laying the groundwork for the introduction of heterogeneous transfer learning as a potential solution. The implementation of heterogeneous transfer learning was described, which involves the utilization of a pre-trained WQI model developed using Bhavani River data to improve predictions for the Bharathapuzha river. The application of heterogeneous transfer learning approach in building enhanced river WQI prediction models was elucidated in this chapter. The overall conclusion of the entire research work with research findings and research contributions will be presented in the next section.

Remarks

The paper titled 'Heterogenous Transfer Learning for Modelling and Forecasting River Water Quality Index' has been communicated to Journal of Cloud Computing ***(in review)***

9. CONCLUSION

The thesis titled 'Temporal Fusion Transformer and Transfer Learning for Time Series River Water Quality Index Prediction Model' demonstrated the research work carried out to predict river water quality. This research work showcases the development of novel prediction models for predicting water quality, utilizing cutting-edge deep learning techniques and transfer learning approaches.

Two distinct datasets have been generated using the physiochemical and seasonal parameters of Bhavani River water for the period January 2016 to December 2020, and two distinct datasets have been developed using water quality data of Bharathapuzha River for the period January 2020 to December 2021.

WQI prediction problem has been formulated as a regression task and WQI prediction models have been built using deep learning architectures such as RNN, LSTM, GRU and the specialized architecture TFT. The implementation is carried out in four phases. First, deep learning algorithms such as RNN, LSTM, and GRU were implemented using WQI-PCA and WQI-SA datasets to build accurate WQI prediction models. Secondly, the TFT was used to build improved WQI prediction model using the WQI-SA dataset. In the third phase, a homogenous transfer learning technique was applied with LSTM based pre-trained WQI model to train the networks RNN, LSTM, GRU independently with limited instances in WQI-BP dataset and an enhanced WQI prediction model has been built. The TFT-based pre-trained model jointly with the WQI-BP dataset was used with homogenous transfer learning and a hybrid WQI prediction model has been built. In the last phase, heterogeneous transfer learning was applied with the TFT based pre-trained model and an extended Bharathapuzha WQI-EBP dataset to build a robust WQI prediction model by training TFT.

Exhaustive experiments have been carried out using python libraries with Tensor flow, Keras and Scikit learn. The performances of various WQI prediction models have been evaluated with metrics such as mean absolute error, mean squared error, root mean squared error, and R2 score and the systematic investigation has been carried out. A thorough performance analysis of the WQI models is performed and the following observations are made.

- The prediction rate of WQI models is increased through representation learning of the time series sequence data with RNN, LSTM, and GRU networks.

- EDA performed on the raw time series data enabled to gain the insights of the data which facilitates data standardization.
- Through feature selection, the association between the pool of predictors and the target variable is strengthened which enables deep neural network architectures RNN, LSTM, and GRU to improve the learning of trends in the data.
- Promising results are produced by WQI prediction models as a result of augmenting seasonal data with regular physicochemical properties for training.
- The WQI-prediction model constructed using the TFT approach achieves encouraging results by leveraging the data handling capabilities of the transformer architecture and capturing the temporal patterns in the time series data.
- It is evident that LSTM and TFT-based WQI predictive models can be efficiently utilized as pre-trained models for transfer learning.
- Enhanced by transfer learning, WQI prediction models demonstrate improved accuracy even when trained on datasets with limited instances.
- The hybrid WQI prediction model developed using TFT and homogenous transfer learning has proved to be the good prediction model for forecasting WQI.
- The heterogeneous transfer learning approach implemented with the extended Bharathapuzha dataset and TFT based pretrained model has produced a robust generalized model for predicting WQI.

From the comprehensive experimental results, it is evident that the said objectives of this research work have been met. The research contributions made in this work are summarized below.

- Water quality prediction problem is formulated as a regression task and a suitable solution is provided using deep learning architectures and transfer learning approaches.
- Sufficient real-time data has been collected for learning the trends and patterns in water quality data and EDA has been performed to understand the statistical distribution of the data.
- Significant features are identified and captured from the data collected from the Bhavani River and Bharathapuzha River and four new time series datasets have been developed.
- Specialized deep learning architectures are used to design and develop the WQI prediction models by training time series data and to forecast water quality.
- TFT is used for constructing the WQI prediction model as TFT enhances the interpretability of time series forecasting by enduring temporal variations, and significant events.

- TFT is integrated with transfer learning approaches both homogeneous and heterogeneous and developed hybrid WQI prediction models by leveraging the merits of both TFT and transfer learning.
- A generalized and holistic framework has been developed using transfer learning to predict water quality index which can be applied to any river in India.

The potential challenge of data representativeness and domain shift can be revisited as a scope of future work and the research can be extended to develop a water quality monitoring system for predicting the river water quality of any geographical area by integrating the generalized WQI prediction model. The advancements in the Internet of Things, could enhance the development of predictive models by harnessing the capabilities of predictive analytics on devices or sensors, which could lead to further improvements in real-time monitoring system at any point of time.