Cross-Station Solar Power Prediction: A Transfer Learning Approach with Deep Learning Models

MSc Research Project
Data Analytics

Savan Kumar Pandith
Student ID: x22182934

School of Computing
National College of Ireland

Supervisor: Catherine Mulwa

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Savan Kumar Pandith…………………………………………………………………
Student ID:	x22182934…………………………………………………………………………….……
Programme:	Data Analytics…………………………………… **Year:** 2023-2024.
Module:	MSc Research Project………………………………………………………………….………
Supervisor:	Dr. Catherine Mulwa………………………………………………….………
Submission Due Date:	12-Aug-2024……………………………………………………………………………….………
Project Title:	Cross-Station Solar Power Prediction: A Transfer Learning Approach with Deep Learning Models
Word Count:	7203 **Page Count** 22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Savan Kumar Pandith ………………………………………..

Date: 12-Aug-2024…………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Cross-Station Solar Power Prediction: A Transfer Learning Approach with Deep Learning Models

Savan Kumar Pandith
x22182934

Abstract

This research investigates the effectiveness of hybrid deep learning models in predicting solar power generation using a transfer learning approach. Three models combining Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Transformer (TF) architectures were developed and evaluated CNN-LSTM-TF, CNN-TF, and LSTM-TF. The models were trained on data from a lower capacity solar station and tested on a higher capacity station and the results showed a good performance by all three models in solar power forecasting, in which CNN-TF models emerged as the top performer with R2 score of 95% and lowest error in MSE of 35.63. The superior performance of CNN-TF models suggests that the combination of extracting features by CNN and the transformer-based attention mechanism has a good capability in capturing complex patterns in solar power generation data when compared to the LSTM-TF model, the CNN-LSTM-TF model performed better, with an R2 score of 92%. Also, the comparison study showed that CNN-TF models performed better than the existing studies.

Keywords: Solar power forecasting, Transfer learning, CNN, LSTM, Transformer- attention mechanism, Hybrid models, Deep learning, Cross-station prediction.

1 Introduction

This research aims to advance the field of solar power generation forecasting by developing and comparing innovative deep learning hybrid models by combining CNN, LSTM, and transformer architectures to improve prediction accuracy. The application of transfer learning techniques will address the challenge of limited data availability at certain solar stations, potentially offering a solution for broader implementation across diverse locations.

Over the years, the usage of electrical energy, has been increasing rapidly for building modern cities, setting up new industries and the growth in population. The rising demand for energy is having a major impact on global warming and to reduce the impact, renewable energy sector has been playing a very important role in reducing the increasing energy demand. Among renewable energy resources, solar energy has been a promising source due to its potential and availability. The sun, the source of solar energy, radiates 1367 W/m2 of solar energy across the atmosphere (Das et al., 2018). Solar energy is one of the most popular and fastest-growing renewable resource which has a low panel cost and high efficiency (International Renewable Energy Agency, 2020), which can help in replacing the traditional fuel-based power plants. Solar power plants make use of photovoltaic (PV) cells to convert sunlight into electricity. PV cells are made of semiconductor materials like silicon, it has a photovoltaic effect, which is the fundamental principle behind solar power generation. These PV cells are connected to inverters, which convert the direct current (DC) electricity produced by solar panels into alternating current (AC) electricity which is compatible with the grid. Integrating solar energy into a grid has many benefits to the economy and the environment by reducing greenhouse gas emission and energy production costs.

1.1 Background and Motivation

According to International Energy Agency (IEA), the installation of global solar PV power capacity could exceed 1700 GW by 2030 (Das et al., 2018). The increasing penetration of solar energy into power grids worldwide requires accurate forecasting of solar power generation to ensure grid stability and optimize energy management strategies (Islam et al., 2023). However, the intermittent and fluctuating nature of solar power has a significant challenge in predicting accurate solar power prediction due to the influenced by meteorological conditions like variability and seasonality, which directly affect its integration into a grid. Accurate forecasting of PV plant generation is essential for the utilities as the deficit between demand and solar generation must be compensated with other energy sources. There are four main types of PV power forecasting techniques like artificial intelligence, hybrid, statistical, and physical techniques. Deep learning models have shown better performance when compared to statistical and machine learning models. Based on experimental results, deep learning-based approaches have been increasingly used in academic research due to their popularity in recent years for forecast solar power (Alkhayat and Mehmood, 2021). Although these approaches have demonstrated potential, they frequently encounter constraints concerning their applicability to diverse geographic regions, climates, and PV system arrangements. Furthermore, a major barrier for solar power generation can be the availability of high-quality, labelled training data, especially for smaller PV systems or those installed in areas with little to no historical data. The need to create solar power forecasting models that are more precise, flexible, and data-efficient to get around these constraints drives this research.

This research is inspired by Saramas et al. (2022), which uses the principle of transfer learning by training the model on low-capacity generation plants and predicting solar power generation on a high-capacity plants. Transfer learning (Li et al., 2020) is a machine learning method that uses information from one domain to improve performance on another similar domain. Through knowledge transfer from an extremely data-rich source PV system to a target PV system with limited information, we may be able to improve the accuracy and stability of solar power predictions even in situations with little data. There is much research conducted related to transfer learning, and one of the papers Ribeiro et al. (2018) considers seasonality and trend factors as an example of the Hephaestus method for cross-building energy forecasting. Similar studies by Fan et al. (2020) have been developed for short-term building energy predictions and energy consumption forecasting with poor-quality data (Gao et al., 2020). However, few studies have addressed the problem of PV production forecasting with transfer learning, allowing room for further research in this area.

As energy demand increases it is important to match real-time demand by building new models that can forecast solar power with high accuracy. Recently attention-based mechanism proposed in Vaswani et al. (2017) has gained popularity, especially in natural language processing, where the transformer model proposed, used language as input data. Transformer-based model was first introduced by Wu et al. (2020) to forecast time-series data, in which self-attention mechanism was used as the strategy to learn complex patterns and dynamics from time series dataset. The transformer-based models have shown improved performance for PV forecasting (Kothona et al., 2022). Additionally, Kim et al. (2021) developed a transformer network to provide accurate solar power forecasting which showed significantly improved results when compared to linear regression, CNN, and LSTM. Even though lot of research had been done in this area, there is still scoped to improve the accuracy of prediction based on other deep learning hybrid models.

1.2 Research Question, Objective and contributions

Based on the gaps and opportunities identified in solar power forecasting, this research goal is to explore the potential of advanced deep learning techniques combined with transfer learning approaches. Specifically, this research seeks to address the following questions:

RQ: *How well can hybrid deep learning models (CNN, LSTM and Transformer) predict solar power generation by using transfer learning approach?*

Sub RQ: *"Can a hybrid model combining CNN-LSTM and transformer architecture outperform other hybrid models in terms of prediction accuracy?"*

The main aim of this research is to develop and compare advanced deep learning hybrid models for solar power generation forecasting, with a focus on transfer learning applications. The specific objectives are:

Obj.1 To collect and preprocess historical solar data from 8 different solar stations, covering the period from January 2019 to December 2022.
Obj.2 Develop three deep learning hybrid models combining the best architectures of CNN, LSTM, and Transformer-based approaches for solar power forecasting.
Obj.3 Implement and evaluate a transfer learning approach, using data from station 8 for training and station 4 for testing, to assess the models' generalization capabilities.
Obj.4 Evaluate the performance of the developed hybrid models using metrics such as Mean squared Error (MSE), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).
Obj.5 Comparison of developed models.
Obj.6 Comparison of developed models with existing modes.

By achieving these research objectives this study introduces a hybrid deep learning model combining CNN, LSTM, and Transformer architectures with transfer learning to enhance solar power forecasting accuracy and provides a thorough evaluation by comparing to existing models. The rest of the technical report is structure as follows. Chapter 2 presents a comprehensive literature review, focusing on deep learning approaches using transfer learning and hybrid models in solar power. Chapter 3 describes about the methodology that is being proposed to solve the research question and sub research question. Subsections about proposed datasets, data collection design process flow, model to be developed and evaluation metrics being considered. Chapter 4 discusses on the design specifications of the developed models. Chapter 5 demonstrates how the models have been implemented. Chapter 6 will discuss the evaluation the results of the implemented models and chapter 7 will be concluded the research project with future work.

2 Literature Review

This chapter provides a critical review of different research papers in the domain of solar power generation and the importance of using transfer learning approach in deep learning. This review is structured into two main subsections:

2.1 A Critical Review of Deep Learning Approaches using Transfer Learning

In recent years deep learning has gathered lot of attention from researchers due to its increased performance when compared to statistical models. The major difference between deep learning and other machine learning models is that deep learning can automatically learn valuable characteristics from dataset. According to Phan et al. 2022 deep learning models based on Recurrent Neural Networks (RNNs), including LSTM, Gated Recurrent Unit (GRUs), and CNNs, are most used architectures. With deep learning models performing well in predicting solar power forecasting, there are still some challenges related to data scarcity and model generalization across different locations and conditions. To address this issue, researchers have started focusing on the concepts of transfer learning. In one of the study, Pratt (1993) developed a discriminability-based transfer (DBT) method between neural networks, which uses information measure to estimate the utility of hyperplanes defined by source weights in the target network and rescales transferred weight magnitudes accordingly. Several experiments shows that target networks initialized via DBT learn significantly faster than networks initialized randomly. Transfer learning idea gained the importance in machine learning in mid -2000s through a survey conducted by Pan and Yang (2010), which helped in defining and categorizing transfer learning approaches. The application of transfer learning in solar power forecasting has evolved significantly in the last decade. In their research, Huang et al. (2017) explored synthetic aperture radar (SAR) target classification with limited labelled data. Although this paper did not focus on solar power generation it provided insights into SAR target classification. Saramas et al. (2022), showed the importance of transfer learning in solar power forecasting for limited data in which they proposed three transfer strategies based on a stack LSTM model. Their transfer knowledge strategy trained the base model from source PV plants to a target PV plant, and the results showed that transfer learning models outperformed the conventional LSTM model. In another study, Zhou et al. (2020) used transfer learning strategies to check the effectiveness of solar power forecasting for a newly built PV plants with limited historical data. They proposed a pre trained LSTM model, and the results showed that transfer learning method reduces prediction errors compared to model trained solely on the limited data. However, the reviewed papers do not directly compare the performance of the transfer learning models with other deep learning hybrid models or explore the impact of different transfer learning techniques. Therefore, this research aims to identify gaps where transfer learning can be utilized with other deep learning transformer-based hybrid models for solar power forecasting. The next chapter discusses the research related to transformer-based hybrid models in deep learning.

2.2 Review of Hybrid Models: CNN-LSTM and Transformer architecture

This chapter reviews how deep learning models like CNN, LSTM and Transformer architectures have contributed to solar power forecasting and research related to hybrid-based models.

CNN models have shown significant promise in solar power forecasting as they help in extracting spatial features from the historical time series data. In one study, Kartini et al. (2022) developed a new unsupervised deep learning CNN model based on weather variable in which they used input data on solar irradiance to improve economic value. The network utilized multiple layers of CNN to extract features from input data, and the results showed an RMSE value of 12.1 W/m2, which is good for short-term solar irradiance prediction. Another study by Hoai Thu et al. (2022) proposed a CNN-LSTM network integrated with the ensemble empirical mode decomposition (EEMD) method to make short-term forecast of solar irradiation in Vietnam. The results showed that the proposed model is better when compared with other single models of CNN, LSTM and Bi-directional-LSTM. However, while CNN models have shown promise, their applications in solar power forecasting are still evolving in hybrid-based models.

LSTM models have always been one the best in solar power forecasting. In one research project, Hari et al. (2022) used a deep learning approach by using LSTM and stacked LSTM for solar irradiance prediction. Their study focused on forecasting direct horizontal irradiance (DHI) using single-layered LSTM and stacked LSTM models with two and three layers. Performance was tested using two datasets from an Indian solar power plant and an Australian solar power plant in which, single-layered and stacked LSTM models showed a better performance. However, this research had limitations as it did not explore other deep learning models like transformers or hybrid approaches.

Transformer is a deep learning model architecture developed by Google, based on the muti-head attention mechanism proposed by Vaswani et al. (2017) in “Attention is All you need.” transformers require less training time when compared to earlier RNN’s as it does not have recurrent units, such as LSTM. The study by Phan et al. (2022), explored the use of transformer-based deep learning model for short-term photovoltaic (PV) power forecasting. They aimed to improve the predictive accuracy by utilizing transformer architectures to identify intricate relationships and patterns in time-series data and the results showed that the transformer-based model performed better than conventional techniques like artificial neural networks (ANNs), (LSTM), and GRUs. In another study, Kim et al. (2021) developed a transformer-based model to improve the prediction accuracy of solar power generation. They adjusted the existing transformer model that was initially created for language translation to solar power forecasting and the results showed that the transformer-based model performed better than conventional techniques like linear regression, 1D-CNN, and LSTM. Also, in another study by Sherozbek et al. (2023) they introduced a transformer-based encoder model for forecasting hourly output for transparent and non-transparent PV systems. Data was collected from Jan to dec 2021 from Buan-gun, Republic of Korea and the performance of the transformer-based model was compared to RNN models, like GRUs and LSTM, and the results showed that transformer-based model performed better than the LSTM and GRU models. For transparent and non-transparent PV modules the transformer-based model recorded a mean absolute error (MAE) of 0.05 kWh and 0.04 kWh, and root mean square errors (RMSE) of 0.24 kWh and 0.21 kWh. The review conducted in earlier papers showed that transformer-based models have a good potential for enhancing PV power forecasting accuracy.

Despite the advancements in transformer models for solar power forecasting, very few researchers have developed hybrid models with a combination of CNN-LSTM and a transformer for solar power forecasting. Al-Ali et al. (2023) and Salman et al. (2024) both explored hybrid deep learning model for this purpose. Al-Ali et al. (2023) proposed a CNN-LSTM-Transformer model using Fingrid open dataset from a solar power plant in Finland, and evaluated the model using RMSE, MAPE and MAE and the model outperformed several baseline models, with the lowest RMSE and MAE values. On the other hand, Salman et al. (2024) investigated various combinations of CNN, LSTM, and transformer models, like LSTM-TF, CNN-LSTM, and CNN-LSTM-TF. They used four-year solar power dataset from France and evaluated their models using MAE, MSE, and RMSE and in the results showed that the CNN-LSTM-TF model particularly with Nadam optimizer, outperformed other models, with the lowest MAE value.

The learning from both literature sections have shown a very good foundation for transfer learning and transformer-based hybrid models in predicting solar power forecasting. However, further research is needed to check how well these models can predict solar power generation using a transfer learning approach. The following chapter will conclude the review of research question and objective 6. The next chapter will discuss the solar power forecasting research methodology.

3 Research Methodology

In today's data-driven world, manufacturing and energy management are just two areas where accurate predictive modelling is crucial for optimizing operations. Predictive forecasting is significantly complicated when dealing with stations or systems that have varying capacities or operating patterns. The issue arises when attempting to predict outputs or performance metrics for higher-capacity stations using historical data from peers with lesser capacity, this issue is very apparent. To address this problem, this study uses past data from a lower capacity station to estimate results for higher capacity station. This approach is useful for direct forecasting, which is a very challenging due to the differences in operational scales and data characteristics between the two stations. This approach will also help in understanding the relationship between the two stations without being constrained by the need for short-term forecasts. This approach can be used in energy management, renewable energy providers, manufacturing plants for predicting the outcomes for large production lines from smaller ones helping in scaling operations.

To implement this approach and to develop a comprehensive predictive model, the research utilizes the 'Cross Industry Standard process for Data Mining' (CRISP -DM) methodology, Figure 1 (Al-Ali et al., 2023) shows a high-level view of the research methodology and how the process has been implemented in this research which is discussed in section 3.1 to 3.3. Also, the design specification of the data collection process is discussed in section 3.2 as illustrated in Figure 2 (Chen et al., 2022).

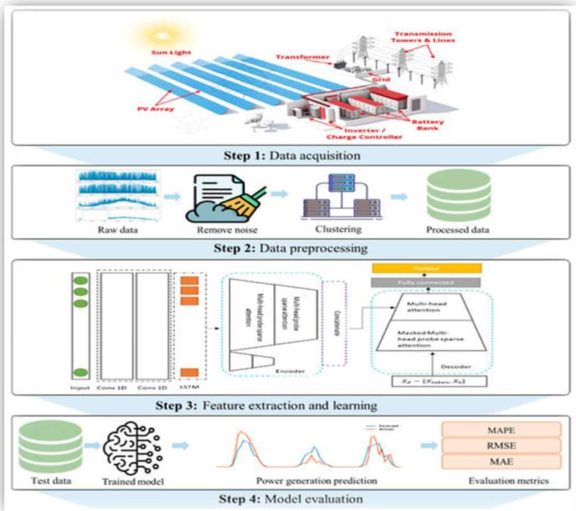

Figure 1: Research Methodology.

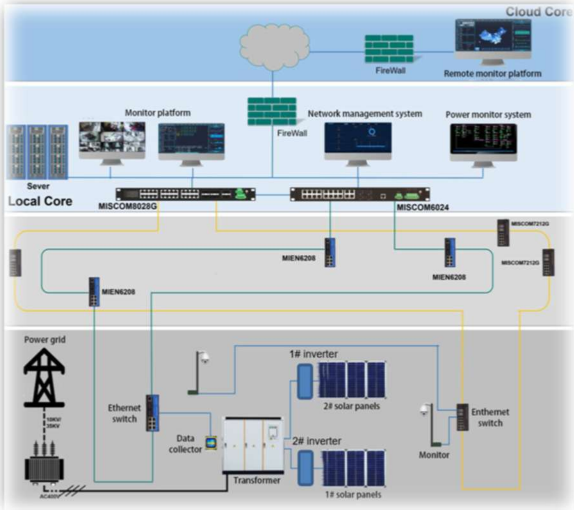

Figure 2: Sensor architecture and data process of the solar stations.

3.1 Data Gathering and Data Understanding

To develop a good robust predictive model, it is very important to capture accurate historical data. To predict an accurate solar power forecasting, the factors utilized are solar irradiation, temperature, wind direction and speed. The data utilized in this research is sourced from a study by Chen et al. (2022), in which the dataset was originally collected from eight solar stations across China and was used for a competition hosted by Chinese state grid in 2021. In this research the study was initially conducted on all 8 solar power stations, but to achieve the research goal and objectives only two solar power stations data was furthered utilized for training and testing based on the available features, lowest nominal capacity and highest nominal capacity across all stations.

Data was collected under the Supervisory Control and Data Acquisition (SCADA) technology, with intervals of every 15 minutes over a period of two years, from Jan 2019 to Dec 2020, as shown in Figure 2. For this research, the data was downloaded in .xlsx format from the repository of the existing study by Chen et al. (2022). Only eight solar power stations dataset was used, that ranged with nominal capacities from 30 MW to 130 MW. Table 1 provides details on the nominal capacity of each station, along with number of observations for each feature.

Table 1: Stations & feature observation count

	Nominal Capacity	Time(year-month-day h:m:s)	Total solar irradiance (W/m2)	Direct normal irradiance (W/m2)	Global horizontal irradiance (W/m2)	Air temp (°C)	Atmosphere (hpa)	Power (MW)	Relative humidity (%)
Station1	50MW	70176	70176	70176	70176	70176	70176	70176	
Station2	130MW	70176	70176	70176	70176	70176	70176	70176	
Station3	30MW	20352	20352	20352	20352		20352	20352	20352
Station4	130MW	70176	70176	70176	70176	70176	70176	70170	70176
Station5	110MW	70176	70176	70176	70176	70176	70176	70176	70176
Station6	35MW	70176	70176	70176	70176	70176	70176	70176	70176
Station7	30MW	70176	70176	70176	70176	70176	70176	70176	70176
Station8	30MW	69408	69408	69408	69408	69408	69408	69408	69408

3.2 Data Cleaning, Preparation, and Exploratory Data Analysis

Data cleaning and preparation process is one of the first important step which lays the groundwork for a detailed understanding of the dataset by addressing the data quality issues. The data was first imported in a Jupyter Notebook using Python libraries for all the solar stations. Based on initial analysis of the total power generated by each solar power station and considering common features (see Table 1), further analysis was conducted only on station 4 and 8.

The next step involved is to clean the data for missing values. The dataset had 6 missing values, which were removed as they contributed to less than 1 % of the total records. Columns were renamed for consistency. Since the data is recorded every 15 minutes over two years, extracting temporal features from the 'Time' column is a crucial step in time series related forecasting as it helps in understanding the underlying temporal structure. 'Time' column for both stations was converted to a datetime format, and features like quarter-hour, hour of the day, day of the week, and month were extracted, as these features help in understanding temporal aspects like intra-day variations, weekly cycles, and seasonal trends.

To improve the model's ability to capture temporal relationships in solar power data, lag features like 'Lag 15min', 'Lag 30min', and 'Lag 1hour' are created by shifting the past values of the 'Power (MW)' variable by 15 minutes, 30 minutes, and 1 hour, respectively. These lag features, along with temporal patterns and short-term fluctuations are considered to help the model to make better predictions by learning from past data. Additionally, features like 'Rolling Mean 1hour' and 'Rolling Std 1hour' are extracted using the mean of power value as they help in smoothing the data, reduce the short-term fluctuations, and capture the longer-term trends and variability. After feature extraction, it was important to check for missing values again as this process led to creation of nine new columns in which less than 1% missing values are observed for some of the features. Dropping these records did not hamper much on the dataset, and further analysis was conducted with station, which had 70,166 records, and station 8, which had 69,044 records.

Exploratory data analysis helps in investigating the underlying patterns and distributions of the data. With the help of different statistical summary and visualization techniques EDA helps in finding important features and detect outliers. To understand the trend and seasonal variation, decomposition chart is created using the statsmodels Python library, in which the output shows a stable trend, and the seasonal component shows a consistent value which indicates a strong yearly seasonal pattern for both solar power stations as illustrated in Figure 3 and Figure 4.

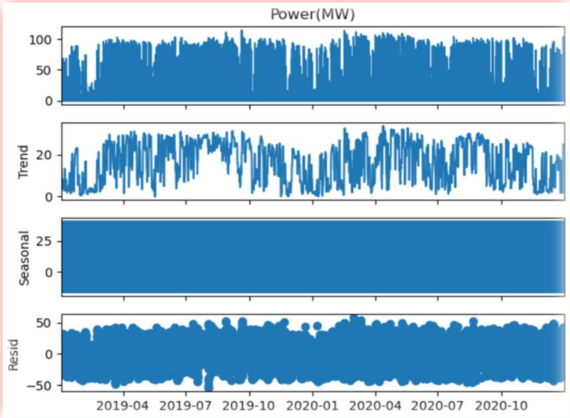

Figure 3: Decomposition chart for station 4.

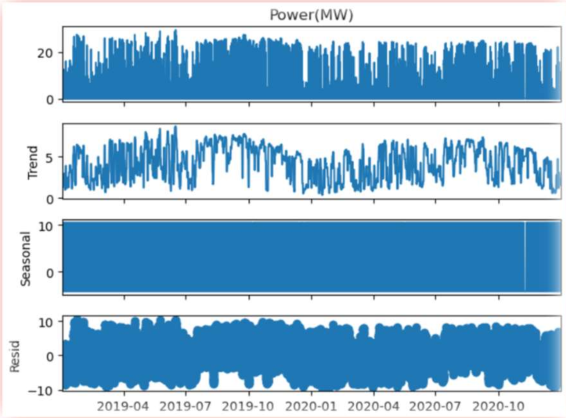

Figure 4: Decomposition chart for station 8.

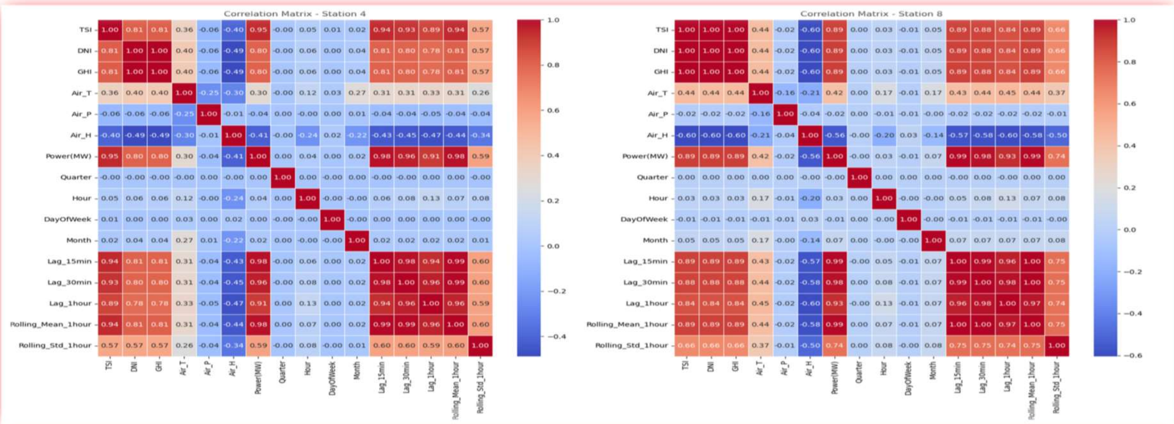

Figure 5: Correlation matrix for station 4 and station 8

Feature selection is performed with the combination of correlation matrix and random forest regressor technique, with Figure 5 showing a heat map of how each feature are correlated. These combined techniques are used in identifying the important feature for model building process. In conclusion the objective 1 outlined in chapter 1, section 1.2 has been successfully implemented at the end of this section. The next section will discuss about different model architecture and evaluation techniques used in this research.

3.3 Data Modelling and Evaluation

This research explores various hybrid models architectures, like CNN LSTM and Transformer, to enhance solar power forecasting. The hybrid approach aims to leverage the strength of each component to improve prediction performance.

Convolutional Neural Networks (CNNs) are among the top deep learning models that have performed very well under different applications. The foundation of a CNN model has four layers i.e. convolution layers, activation layers, pooling layers, and fully connected layers as shown in Figure 6 (Al-Ali et al., 2023). The operation of a convolutional layer is to extract the spatial features in which the activation layers enhance the learning capability by maximizing the non-linearity which allow the network to learn complex patterns. Pooling layers help compress the dimension of the feature maps and the fully connected layers combines all the features extracted by the convolution layers.

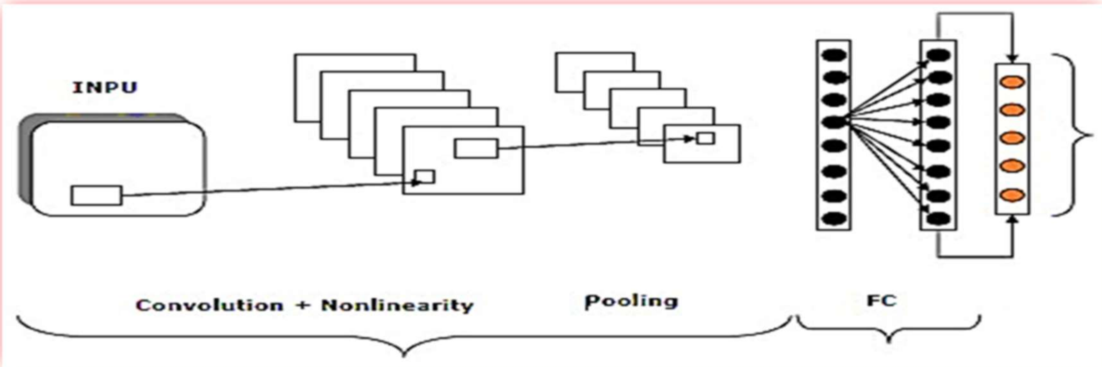

Figure 6: CNN architecture

Long Short-Term Memory (LSTM) is designed to overcome the limitations of traditional RNNs. LSTMs maintain and manipulate information over extended periods and suitable for sequence and time series data. LSTM was originally developed to address the vanishing gradient problem which encountered in traditional RNNs. The key components of the LSTM

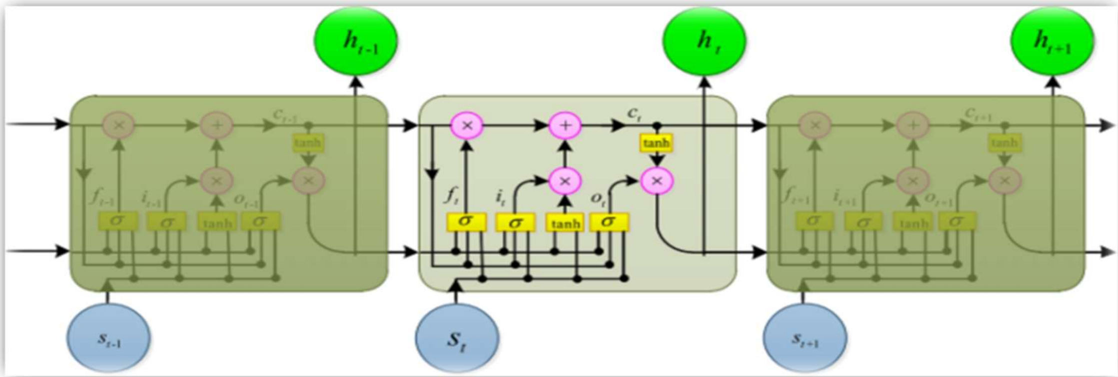

Figure 7: LSTM architecture

architecture is made of a memory cell, input gate, forget gate, output gate and a cell state as shown in Figure 7 (Salman et al., 2024). The process starts as information first flows into the LSTM unit, where the forget gate checks what to discard from the previous cell state. Next, the input gate checks what new information should be stored. The cell state is then updated by forgetting irrelevant data and updating new information. Next, the output gate selects what information is needed from the updated cell state, which is passed to the next layer. This process is continuously repeated for each time step in the sequence, which makes the LSTM to maintain relevant information over time while removing unnecessary data.

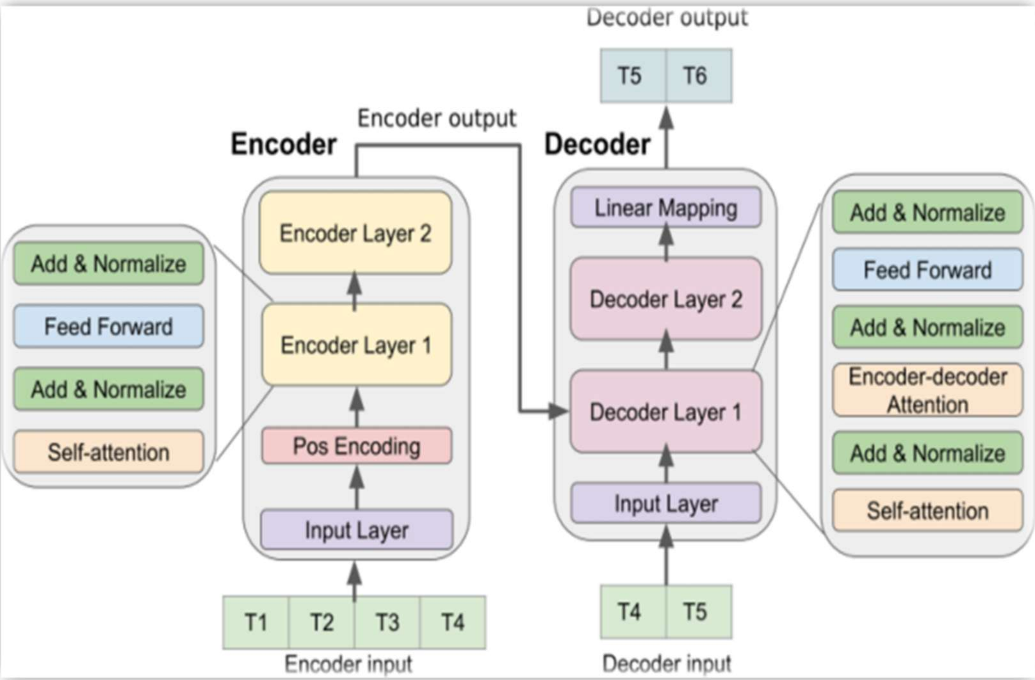

Figure 8: Transformer architecture

Transformer model was first introduced in 2017 by Veltman et al., specifically for natural language processing. Its usage has been increasing in various fields, including solar power forecasting. Transformer architecture begins with input embedding and positional encoding, which helps model to understand both the content and order of input sequences. The heart of the transformer as shown in Figure 8 (Salman et al., 2024), lies in its multi-head attention mechanism, which makes the model focus on different parts of the input simultaneously. The attention is applied in both self-attention layers, where the model relates different positions of a single sequence, and in encoder-decoder attention layers, where the model can attend to relevant parts of the input sequence for each output element. The model then processes this attended information through feed-forward neural networks. In the decoder, an additional attention layer is used that attends to the encoder's output and finally, the model gives output through a linear layer.

When compared to previous recurrent models, Transformers have a good ability to capture long-range dependencies and the sequence process running in parallel makes it valuable for solar power forecasting. Transformers have capability to handle complex temporal patterns and multiple influential factors in solar energy prediction, and it can also improve the accuracy of solar power predictions by considering different features, like historical power output, weather forecasts, and data from satellites. With these improved forecasting capabilities of transformer model, it is beneficial for grid operators and energy dealers, to integrate the solar power into electricity networks and allows a global shift to renewable energy sources.

Evaluation metrics are the measures used to assess the performance of predictive models. Metrics are important as they help in quantify model accuracy, allow comparison between different models, help in model selection and improvements. They also provide valuable insights into model reliability and generalizability. The key metrics used in this research to check the model evaluations are:
Mean Absolute Error (MAE): measures the average absolute difference between predicted and actual values.
Mean Square Error (MSE): calculates the average of squared differences between predicted and actual values.
Root Mean Square Error (RMSE): the square root of MSE, providing a measure in the same units as the target variable.

These methodological stages are very essential steps in building a robust model, that helps in managing the difficulties associated with predicting solar power output that concludes Chapter 3. The design specifications of the predictive models, including detailed configurations is covered in the following Chapter 4.

4 Design Specification

This chapter gives a detailed overview of a high-level design of transfer learning approach and different the hybrid deep learning models designed for solar power prediction by combining convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and Transformer based multi-head attention mechanisms.

4.1 Transfer learning method

Transfer learning is a technique that uses source domain to enhance the performance of a target domain. The process starts with model training on a base problem as shown in Figure 9 (Saramas et al.,2022).

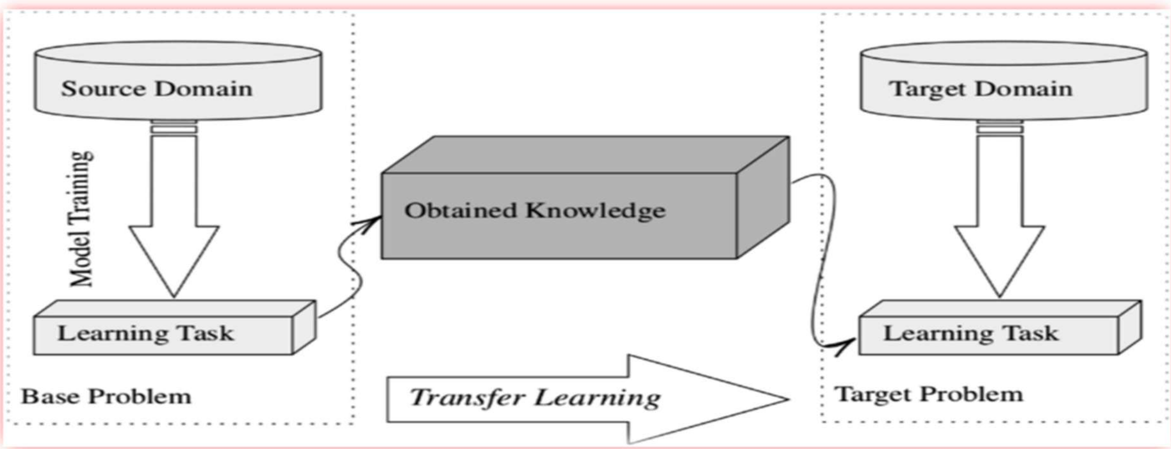

Figure 9: Transfer Learning method

This pre-trained model, built in the source domain, captures essential features and patterns and these insights are then applied to a different but similar target problem. The architecture of transfer learning has two primary components, the target domain, which applies the knowledge that has been transferred, and the source domain, where the initial learning took place. This method allows the model to start from a better educated state, potentially improving prediction accuracy and model robustness. It is especially useful when the target domain has limited data.

4.2 CNN-LSTM-TF Hybrid Model Design

The CNN-LSTM-TF hybrid model shown in Figure 10, is inspired by the work of Al-Ali et al. (2023) and consists of a two 1D-CNN layer, one LSTM layer and a multihead attention-based transformer. Each component plays a distinct role in extracting important features based on its architecture. The CNN layer helps in extracting the spatial features, the LSTM later extracts the temporal features, and the transformer employs the extracted features to generate the forecasting results. The transformer's encoder–decoder strategy has the potential of improving forecasting accuracy by learning from the mixed spatial and temporal features.

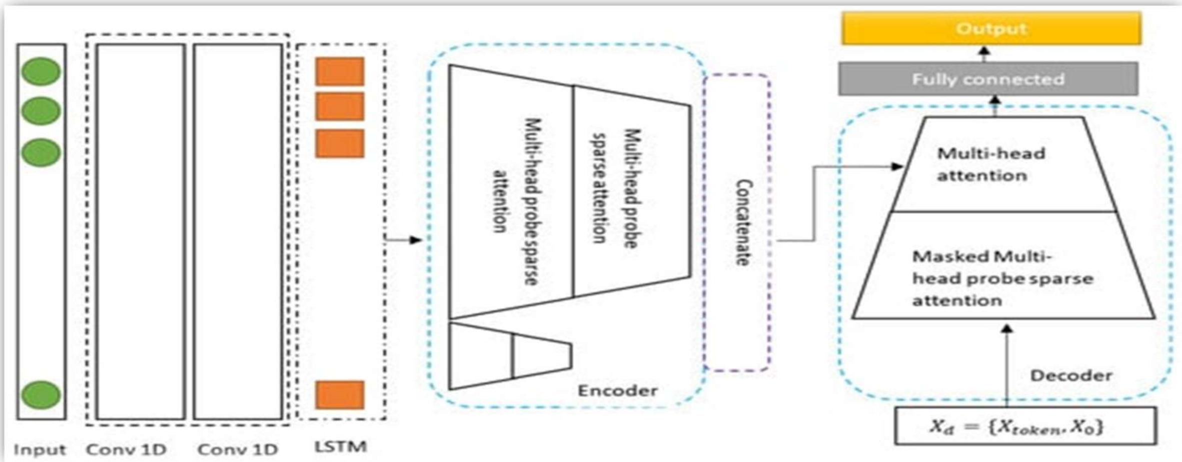

Figure 10: CNN-LSTM-TF Model

The Typical architecture of the CNN-LSTM-TF model as shown in Figure 10, starts with an input layer which accepts the time series data and the selected features that were discussed in section 3.2. The input shape is made flexible that captures various temporal resolutions and feature sets that are common in solar forecasting. The input layer sends the data to the two 1D convolutional layers in which both layers use 64 filters with a kernel size of 3 and same padding to ensure that output size matches the input size also ReLU activation functions is used to help with non-linearity. These layers are effective in capturing local patterns and short-term dependencies in the input data, which is particularly useful for identifying weather-related patterns that influence solar power generation. Following the CNN layer, a single LSTM layer with 64 units processes the data, returning sequences which helps capturing long-term temporal dependencies and seasonal patterns. The processed data is then sent to the custom multi head attention layer which allows the model to focus on different parts of the input sequence. The attention mechanism helps the model identify and weigh the most relevant features and time steps for prediction, which is crucial in handling the variability of solar power due to factors like cloud cover and atmospheric conditions. Decoder structure concatenates the LSTM output with the attention output and applies another attention layer to this concatenated output. This structure helps the model to focus on the most important features. In the end the output dense layer with linear activation produces the final prediction, which is suitable for the regression task of forecasting solar power output. In addition to the CNN-LSTM-TF model, other variants like CNN-TF and LSTM-TF are designed with the same principal.

These hybrid deep learning models combined with advanced neural network architectures can capture both short-term and long-term patterns, along with the attention mechanism's focused on relevant features, makes it a promising approach to address the complexities and helps in accurate solar power prediction. This concludes the model design chapter, with following chapter 5 discussing about how the model has been implemented.

5 Implementation of Cross-Station Solar Prediction Models

This implementation section discusses the detail view of how previously described methodologies and model designs for solar power forecasting are put into practice. After selecting the important features as discussed in section 3.2, it is important to prepare the data as per the model requirements.

All the three models were trained on station 8 and tested on station 4. The dataset was first isolated by separating the input features from target variable (Power) for both station 4 and 8. Min-Max scalar function used from 'sklearn' Python library to normalize both the stations. The scaler fit is applied to the feature and target data of station 8 to ensure consistent scaling as shown in Figure 11. Also, this step is crucial in this research as the models are trained only on station 8 dataset and then evaluated on station 4, which also relates back to the research objective 3 and research question. Next the data for station 8 is split into train (80%) and validation (20%) by using 'train test split' function from 'sklearn' Python library.

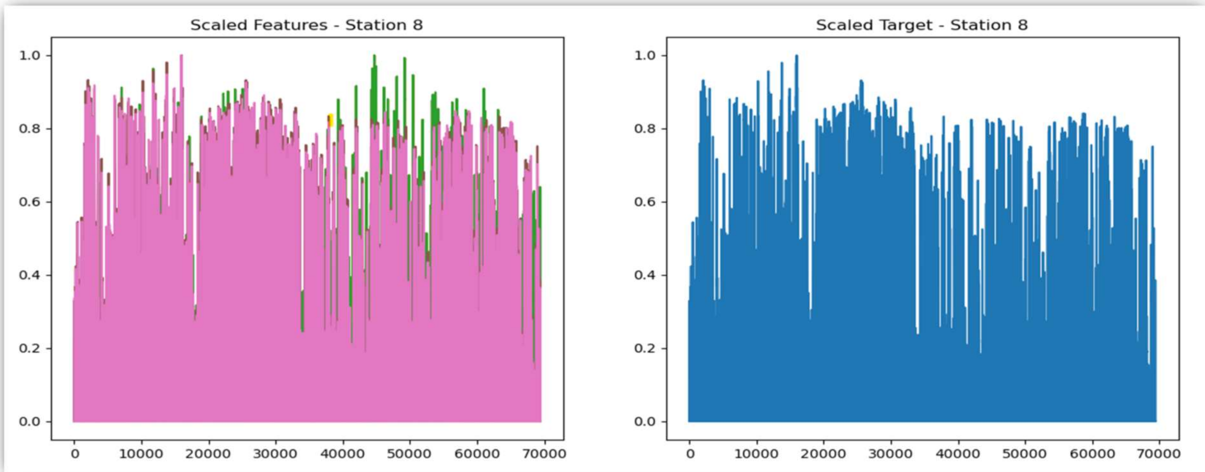

Figure 11: Scaled Features & Target for station 4 and 8

5.1 Implementation of CNN-LSTM-TF Model

The first model implemented in this research is the CNN-LSTM-TF hybrid model. Figures 12 and 13 shows detailed summary of the model's architecture and the loss plot during the training process. For model training, the MSE loss function is used with 50 epochs and an early callback is set to 15 allowing model to train properly and to stop if the validation loss does not improve after 15 epochs. The input layer of the model first takes the scaled dataset which is fed into two convolutional layers(conv1D), which helps in extracting local features from the input sequence. These features are then fed into LSTM layer to capture long term dependencies in the data. The

Layer (type)	Output Shape	Param #	Connected to
input_layer_3 (InputLayer)	(None, 7, 1)	0	-
conv1d_6 (Conv1D)	(None, 7, 64)	256	input_layer_3[0]…
conv1d_7 (Conv1D)	(None, 7, 64)	12,352	conv1d_6[0][0]
lstm_1 (LSTM)	(None, 7, 64)	33,024	conv1d_7[0][0]
multi_head_attenti… (MultiHeadAttentio…	(None, None, 64)	16,640	lstm_1[0][0], lstm_1[0][0], lstm_1[0][0], dense_22[0][0], dense_22[0][0], dense_22[0][0]
concatenate_3 (Concatenate)	(None, 7, 128)	0	lstm_1[0][0], multi_head_atten…
dense_22 (Dense)	(None, 7, 64)	8,256	concatenate_3[0]…
get_item_3 (GetItem)	(None, 64)	0	multi_head_atten…
dense_23 (Dense)	(None, 1)	65	get_item_3[0][0]

Total params: 70,593 (275.75 KB)
Trainable params: 70,593 (275.75 KB)
Non-trainable params: 0 (0.00 B)

Figure 12: CNN-LSTM-TF Model Summary

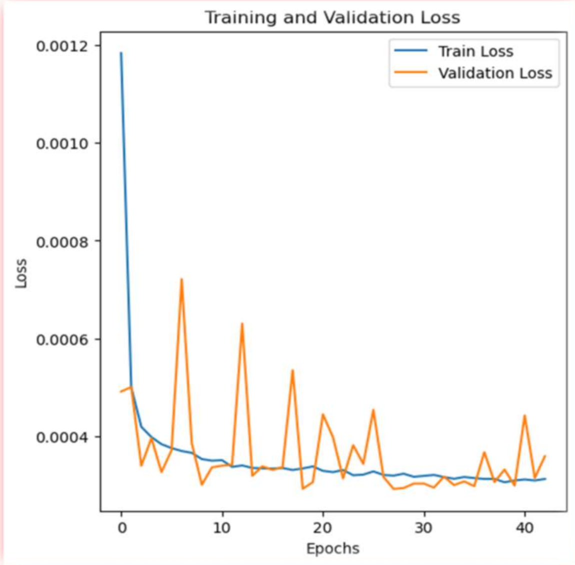

Figure 13: CNN-LSTM-TF Train & Val loss

output of the LSTM layer is fed into the multi-head attention layer, which allows the model to focus on different input sequence simultaneously and capture complex relationships in the data. The output from the attention layer is concatenated with LSTM output which gives the combination of both sequential and attention- based features. Finally, 2 dense layers are added to reduce the dimensionality to produce the final output which contained a total of 70,693 trainable parameters. Model completed only 43 epochs as shown in Figure 13, training loss decreases sharply in the first few epochs which indicates model quickly learned to fit the training data. But the validation loss shows some fluctuations in the early stages, which can be due to either sensitivity or complexity of the data. By the end of the training both training and validation loss stabilized around 0.0003- 0.0004 indicating a good balance fit.

5.2 Implementation of CNN-TF Model

CNN-TF is the second model implemented with the detailed architecture summary shown in Figure 14. & the corresponding loss plot in Figure 15.

Layer (type)	Output Shape	Param #	Connected to
input_layer_1 (InputLayer)	(None, 7, 1)	0	-
conv1d_2 (Conv1D)	(None, 7, 64)	256	input_layer_1[0][0]
conv1d_3 (Conv1D)	(None, 7, 64)	12,352	conv1d_2[0][0]
multi_head_attention_1 (MultiHeadAttention)	(None, None, 64)	16,640	conv1d_3[0][0], conv1d_3[0][0], conv1d_3[0][0], dense_10[0][0], dense_10[0][0], dense_10[0][0]
concatenate_1 (Concatenate)	(None, 7, 128)	0	conv1d_3[0][0], multi_head_attention_1[0]…
dense_10 (Dense)	(None, 7, 64)	8,256	concatenate_1[0][0]
get_item_1 (GetItem)	(None, 64)	0	multi_head_attention_1[1]…
dense_11 (Dense)	(None, 1)	65	get_item_1[0][0]

Total params: 37,569 (146.75 KB)
Trainable params: 37,569 (146.75 KB)
Non-trainable params: 0 (0.00 B)

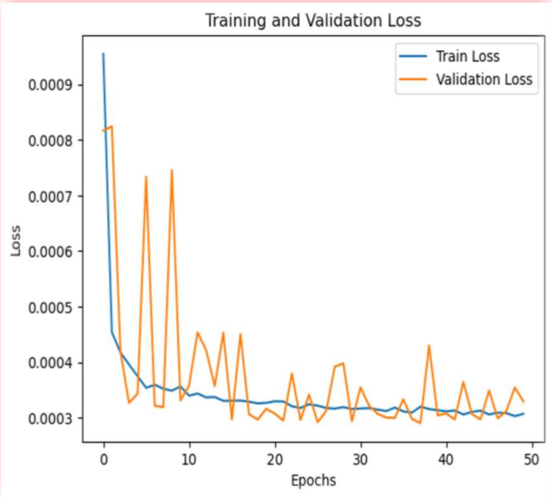

Figure 14: CNN-TF Model Summary Figure 15: CNN-TF Train & Val loss

This model has a very similar to architecture and parameters used as discussed in the previous CNN-LSTM-TF model, except the removal of LSTM layer. CNN-TF model contains 37,569 trainable parameters and it completed all 50 epochs as shown in Fig. 13. Training loss shows a similar dip as the CNN-LSTM-TF model, but validation loss patterns were high initially and follows a downward trend in the end. Both training and validation loss stabilized around 0.0003 indicating a good balance fit and slightly better than CNN-TF.

5.3 Implementation LSTM-TF Model

LSTM-TF is the third and last model implemented with the detailed architecture summary shown in Figure 16, with the corresponding loss plot in Figure 17. This model also has a very similar architecture and parameters used as discussed in the previous CNN-LSTM-TF model, except the removal CNN layer and adding two LSTM layers. LSTM-TF model has 74,881 trainable parameters and it completed 45 epochs as shown in Figure 17. During training the loss shows a rapid decrease in both training and validation loss during the initial epochs. In the entire training process validation loss closely follows the training and gradually stabilized around 0.0004 indicating a good balance fit and slightly better than CNN-LSTM-TF.

Layer (type)	Output Shape	Param #	Connected to
input_layer_2 (InputLayer)	(None, 7, 1)	0	-
lstm_1 (LSTM)	(None, 7, 64)	16,896	input_layer_2[0][0]
lstm_2 (LSTM)	(None, 7, 64)	33,024	lstm_1[0][0]
multi_head_attention_2 (MultiHeadAttention)	(None, None, 64)	16,640	lstm_2[0][0], lstm_2[0][0], lstm_2[0][0], dense_16[0][0], dense_16[0][0], dense_16[0][0]
concatenate_2 (Concatenate)	(None, 7, 128)	0	lstm_2[0][0], multi_head_attention_2[0]…
dense_16 (Dense)	(None, 7, 64)	8,256	concatenate_2[0][0]
get_item_2 (GetItem)	(None, 64)	0	multi_head_attention_2[1]…
dense_17 (Dense)	(None, 1)	65	get_item_2[0][0]

Total params: 74,881 (292.50 KB)
Trainable params: 74,881 (292.50 KB)
Non-trainable params: 0 (0.00 B)

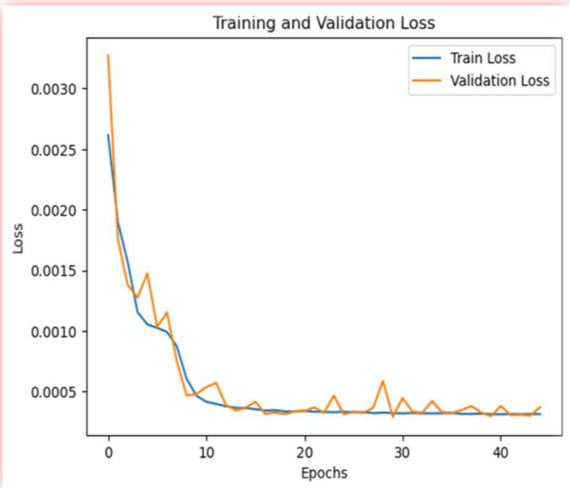

Figure 16: LSTM-TF Model Summary Figure 17: LSTM-TF Model Summary

All three models CNN-LSTM-TF, CNN-TF, and LSTM-TF were trained and then validated, to ensure that there was a suitable balance between fitting the training data and prevent from overfitting. Each model followed the training procedure, loss functions, and total number of trainable parameters. All three models were able to achieve minimal and stable validation losses, indicating that they are very useful in forecasting solar power generation. In conclusion with this chapter the objectives 1, 2 and 3 outlined in chapter 1, section 1.2 has been successfully implemented and the research question posed in Chapter 1, Sub-section 1.2 has been partially addressed in chapter 5. The following chapter 6 will delve deeper into the results, evaluating the effectiveness of these models on test data which will help understand model's performance over unseen data and complete the remaining objectives.

6 Evaluation and Results

This section discusses the evaluation method and the results obtained for all the hybrid models. Evaluation metrics used in this research are MSE, RMSE and MAE and Table 2 shows the evaluated results for all three models.

Table 2: Evaluation Metrics

Model	MSE	RMSE	MAE	R^2 Score
CNN-LSTM-TF	59.168841	7.692128	3.637734	92%
CNN-TF	35.627928	5.968913	3.027626	95%
LSTM-TF	109.75549	10.47643	4.802842	85%

6.1 Evaluation and Results of CNN-LSTM-TF Model

CNN-LSTM-TF model shows a good predictive performance for solar power prediction, with R^2 score of 92%, which indicates that model can effectively explain 92% of the variance in the output. Also, the model shows a low error rates with the MSE of 59.17, RMSE of 7.69, and MAE of 3.64 as shown in Table 2. The plot of actual vs predicted values shown in Figure 18, indicates model's effectiveness in capturing complex patterns and avoids overfitting in solar power forecasting.

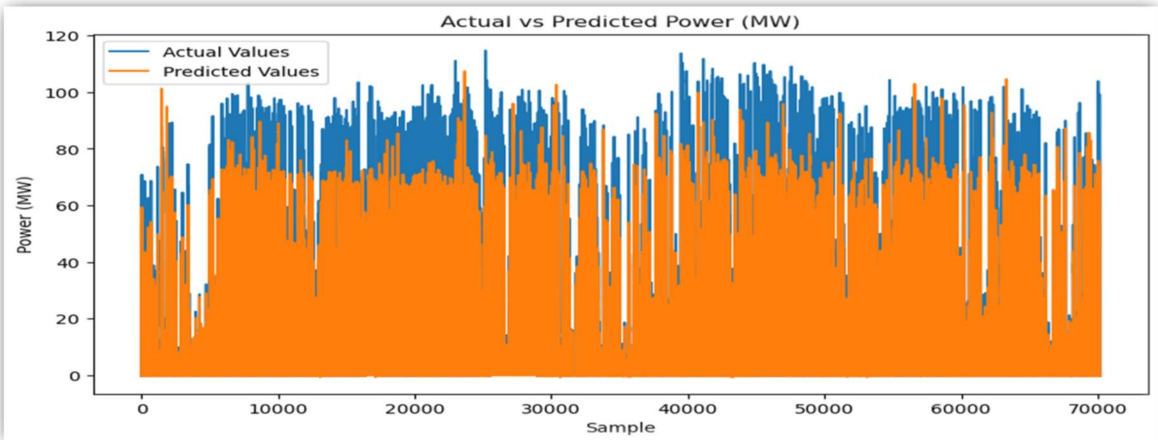

Figure 18: CNN-LSTM-TF Actual vs Predicted

6.2 Evaluation and Results of CNN-TF and LSTM-TF Model

CNN-TF model shows a lowest error rates with MSE of 35.63, RMSE of 5.97, MAE of 3.03 and has the highest R^2 score of 95%. The plot of actual vs predicted values as seen in Fig. 17, indicates that model is fitting better than the other two models by capturing complex patterns of solar power data.
LSTM-TF model has the highest error rates with MSE of 109.76, RMSE 10.48, MAE 4.80 and with the lowest R^2 score of 85%. While the model showed a stable training behaviour and explained significant portion in the variance, but it performed the poorest in evaluation metrics when compared with other two models.

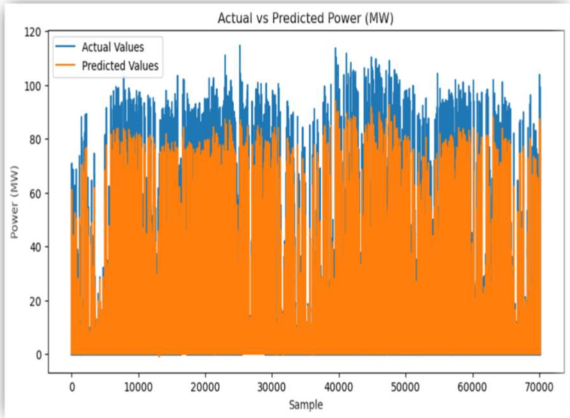

Figure 19: CNN-TF Actual vs Predicted

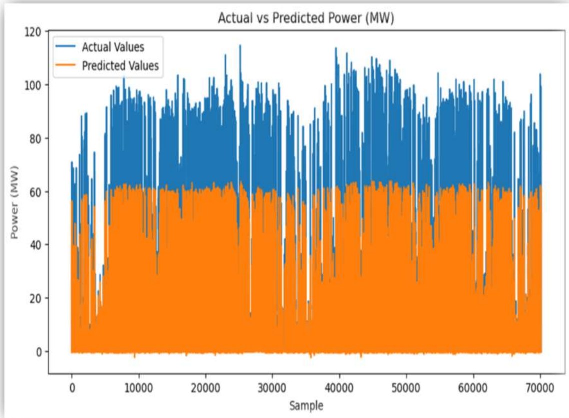

Figure 20: LSTM-TF Actual vs Predicted

In conclusion with chapter 6.2 the research question and the objectives 3 and 4 outlined in chapter 1, section 1.2 has been completed successfully by evaluating and testing all three hybrid models using cross-station approach. The results showed that all three models can predict good solar power prediction using cross-station approach. Next sub section discusses the comparison of these models.

6.3 Discussion and Comparison of Developed Models Verses Existing

This research is an experimental study of using hybrid deep learning models in predicting solar power forecasting using a transfer learning approach and focuses on building three hybrid models CNN-LSTM-TF, CNN-TF, and LSTM-TF, where the models are trained on the data from a lower capacity solar station and then tested on a higher capacity solar station and the all the three models show a good performance in predicting solar power forecasting.

6.3.1 Comparison of Developed Models

CNN-TF model performs the best with a very high R2 score of 95% and has the lowest MSE of 35.63, which shows that the combination of extracting features by CNN and the transformer-based attention mechanism has a good capability in capturing complex patterns in solar power generation data. The CNN-LSTM-TF model performed better when compared with LSTM-TF model, the evaluation metrics showed an R2 score of 92% for CNN-LSTM-TF model, second best after CNN-TF model. LSTM-TF model performed the lowest when compared with other two models, which shows that even though LSTM layers can capture temporal dependencies, it’s not necessary that it can help in improving the performance, because of LSTM addition. Also, when the convolutional layer is removed from the LSTM-TF model the model performance shows the lowest with R2 score of 85%, this is due to the impact of model not extracting spatial features from the input data. These results show that CNN-TF model has a good efficiency in transfer learning method because of the convolutional layers help in detecting the spatial features and when combined with attention mechanisms to detect the temporal patterns in solar power data. These performance results show that, it is very important to choose a right model architecture that can help generalize well from a lower capacity station to a higher capacity by applying transfer learning approach for solar power forecasting. With these finding in chapter 6.3.1, the research sub question outlined in chapter 1, section 1.2 and the objective 5 outlined in chapter 1, section 1.2 has been completed successfully by comparison analysis of developed models and confirming that CNN-LSTM-TF model does not outperform other developed models, but the model does have a good prediction accuracy and is the next best model after CNN-TF.

6.3.2 Comparison of Developed Models vs Existing Models

Table 3 shows the summary of the performance metrics (RMSE and MAE) of the developed models against the models developed in existing studies.

Table 3: Comparison of Developed Models with Existing Models

Model	RMSE	MAE
Developed CNN-LSTM-TF	7.69	3.64
Developed CNN-TF	5.97	3.03
Developed LSTM-TF	10.48	4.8
CNN-LSTM-TF (Al-Ali et al., 2023)	0.344	0.393
CNN-LSTM-TF (Salman et al., 2023)	1.039	0.562
CNN-TF (Salman et al., 2023)	18.98	15.58
LSTM-TF (Salman et al., 2023)	1.106	0.632

The developed CNN-TF model showed a lower RMSE (5.97) and MAE (3.03) when compared to the other existing models as shown in Table 3, but the CNN-LSTM-TF models from the existing studies by Al-Ali et al. (2023) and Salman et al. (2023), performed better in terms of RMSE and MAE. The existing models showed a low error metrics, with CNN-LSTM-TF model by Al-Ali et al. (2023) scored an RMSE of 0.344 and an MAE of 0.393. The developed LSTM-TF model showed the highest RMSE of 10.48 and MAE of 4.80 when compared with existing models, which shows that the model is not as effective as CNN in capturing the complex patterns of solar power generation. While the developed models have shown good performance in predicting solar power, they fall short when compared with other models from

Al-Ali et al. (2023) and Salman et al. (2023), in terms of RMSE and MAE. The existing models, particularly the CNN-LSTM-TF architectures, showed much lower error rates compared to the developed models. This shows that there is lot of room for improvement in the developed models by further optimizing the architecture or tuning hyperparameters. With the findings in chapter 6.3.2, the research objective 6 outlined in chapter 1, section 1.2 has been successfully completed by doing comparison analysis of the developed models with exiting models.

7 Conclusion and Future Work

This research explores the potential of solar power prediction across different capacity stations using advanced hybrid deep learning techniques and transfer learning. To address the research question and sub-question discussed in chapter 1, section 1.2 the main objective was to develop three hybrid deep learning models CNN-LSTM-TF, CNN-TF, and LSTM-TF and train them on data from a lower capacity solar station and tested on a higher capacity solar station. The developed models show a good overall performance in predicting solar power prediction, in which CNN-TF model showed the best performance by achieving 95% accuracy. These results show that transfer learning has lot of potential in solar power forecasting, particularly for new or expanding installations with limited historical data. But on the other hand, the most complex model CNN-LSTM-TF did not outperform the simpler CNN-TF, which suggests that sometimes, less is more in AI design.
However, the research has some limitations as the developed models show a very high error rates when compared to existing models discussed in the section 6.3, particularly by Al-Ali et al. (2023) and Salman et al. (2023), which suggests that there is still room for improvement. For future work the study can include more solar stations with varying capacities and geographical locations to test the robustness and generalizability of the transfer learning approach.

Acknowledgment

I would like to take this opportunity to thank my family for encouraging me in taking this new learning curve and giving their continuous support during this time. I also want to thank my supervisor, Dr. Catherine Mulwa for supporting me in this research. Her continuous weekly sessions and feedback helped me a lot in completing the research project.

References

Al-Ali, E.M., Hajji, Y., Said, Y., Hleili, M., Alanzi, A.M., Laatar, A.H., Atri, M. (2023) 'Solar energy production forecasting based on a hybrid CNN-LSTM-Transformer model', *Mathematics*, 11, p. 676. doi: 10.3390/math11030676.

Alkhayat, G. and Mehmood, R. (2021) 'A review and taxonomy of wind and solar energy forecasting methods based on deep learning', *Energy AI*, 4, p. 100060.

Chen, Y. and Xu, J. (2022) 'Solar and wind power data from the Chinese State Grid Renewable Energy Generation Forecasting Competition', *Scientific Data*, 9, p. 577. doi: 10.1038/s41597-022-01696-6.

Das, U., Tey, K., Seyedmahmoudian, M., Mekhilef, S., Idris, M., Van Deventer, W., et al. (2018) 'Forecasting of photovoltaic power generation and model optimization: A review', *Renewable and Sustainable Energy Reviews*, 81, pp. 912-928.

Fan, C., et al. (2020) 'Statistical investigations of transfer learning-based methodology for short-term building energy predictions', *Applied Energy*, 262, p. 114499.

Gao, Y., Ruan, Y., Fang, C. and Yin, S. (2020) 'Deep learning and transfer learning models of energy consumption forecasting for a building with poor information data', *Energy and Buildings*, 223, p. 110156.

Hari, N. G., & G., J. (2022). Solar irradiance prediction using deep learning-based approaches. *Proceedings of the 2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 1-5. IEEE. https://doi.org/10.1109/CSDE56538.2022.10089282

Hoai Thu, N. T., Van, P. N., Bao, P. Q., Nhat Nam, N. V., Minh, P. H., & Quang, T. N. (2022). Short-term forecasting of solar radiation using a hybrid model of CNN-LSTM integrated with EEMD. *Proceedings of the 2022 6th International Conference on Green Technology and Sustainable Development (GTSD)*, 854-859. IEEE. https://doi.org/10.1109/GTSD54989.2022.9988761

Huang, Z., Pan, Z. and Lei, B. (2017) 'Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data', *Remote Sensing*, 9, p. 907.

International Renewable Energy Agency (IRENA). (2020) *Renewable power generation costs in 2019*. Available at: https://www.irena.org/publications/2020/Jun/Renewable-Power-Costs-in-2019 (Accessed: 15 December 2022).

Islam, M., Shuvo, S.S., Shohan, J.A. and Faruque, O. (2023) 'Forecasting of PV plant output using interpretable temporal fusion transformer model', in *2023 North American Power Symposium (NAPS)*, Asheville, NC, USA, pp. 01-06. doi: 10.1109/NAPS58826.2023.10318698.

Kartini, U.T., Hariyati, W., Aribowo, W. and Wardani, A.L. (2022) 'Development hybrid model deep learning neural network (DL-NN) for probabilistic forecasting solar irradiance on solar cells to improve economics value added', in *2022 Fifth International Conference on Vocational Education and Electrical Engineering (ICVEE)*, Surabaya, Indonesia, pp. 151-156. doi: 10.1109/ICVEE57061.2022.9930352.

Kim, N., Lee, H., Lee, J. and Lee, B. (2021) 'Transformer based prediction method for solar power generation data', in *2021 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju Island, Republic of Korea, pp. 7-9. doi: 10.1109/ICTC52510.2021.9620897.

Kothona, D., Spyropoulos, K., Valelis, C., Sarigiannidis, C., Chatzisavvas, K. and Christoforidis, G. (2022) 'Efficient 24-hour ahead PV energy production forecasting employing a transformer-based model', in *2022 2nd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)*, pp. 1-6.

Li, Y., He, K., Xu, D. and Luo, D. (2020) 'A transfer learning method using speech data as the source domain for micro-Doppler classification tasks', *Knowledge-Based Systems*, 209, p. 106449. doi: 10.1016/j.knosys.2020.106449.

Pan, S.J. and Yang, Q. (2010) 'A survey on transfer learning', *IEEE Transactions on Knowledge and Data Engineering*, 22(10), pp. 1345-1359. doi: 10.1109/TKDE.2009.191.

Phan, Q.-T., Wu, Y.-K. and Phan, Q.-D. (2022) 'An approach using transformer-based model for short-term PV generation forecasting', in *2022 8th International Conference on Applied System Innovation (ICASI)*, Nantou, Taiwan, pp. 17-20. doi: 10.1109/ICASI55125.2022.9774491.

Pratt, L.Y. (1993) 'Discriminability-based transfer between neural networks', in *Advances in Neural Information Processing Systems*, 5.

Ribeiro, M., Grolinger, K., ElYamany, H.F., Higashino, W.A. and Capretz, M.A. (2018) 'Transfer learning with seasonal and trend adjustment for cross-building energy forecasting', *Energy and Buildings*, 165, pp. 352–363.

Salman, D., Direkoglu, C., Kusaf, M., et al. (2024) 'Hybrid deep learning models for time series forecasting of solar power', *Neural Computing and Applications*, 36, pp. 9095–9112. doi: 10.1007/s00521-024-09558-5.

Sarmas, E., Dimitropoulos, N., Marinakis, V., et al. (2022) 'Transfer learning strategies for solar power forecasting under data scarcity', *Scientific Reports*, 12, p. 14643. doi: 10.1038/s41598-022-18516-x.

Sherozbek, J., Park, J., Akhtar, M.S. and Yang, O.-B. (2023) 'Transformers-based encoder model for forecasting hourly power output of transparent photovoltaic module systems', *Energies*, 16, p. 1353. doi: 10.3390/en16031353.

Vaswani, A., Shazeer, N.M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L. and Polosukhin, I. (2017) 'Attention is all you need', in *Advances in Neural Information Processing Systems*, pp. 5998-6008.

Veltman, A., Pulle, D.W.J. and De Doncker, R.W. (2017) 'Attention is all you need', *Power Systems*, 30, pp. 47-82.

Wu, N., Green, B., Ben, X. and O'Banion, S. (2020) 'Deep transformer models for time series forecasting: The influenza prevalence case', *Neural Computation*.

Zhou, S., Zhou, L., Mao, M. and Xi, X. (2020) 'Transfer learning for photovoltaic power forecasting with long short-term memory neural network', in *2020 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 125-132. doi: 10.1109/BigComp48618.2020.00-87.