Enhancing Wave Energy Conversion Efficiency through Deep Learning-Based Forecasting on Ireland's West Coast

MSc Research Project
Msc Data Analytics

Sabarinathan Chandramohan
Student ID: x22213333

School of Computing
National College of Ireland

Supervisor: **Teerath Kumar Menghwar**

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Sabarinathan Chandramohan
Student ID:	x22213333
Programme:	Msc Data Analytics
Year:	2024
Module:	MSc Research Project
Supervisor:	**Teerath Kumar Menghwar**
Submission Due Date:	16/09/2024
Project Title:	**Enhancing Wave Energy Conversion Efficiency through Deep Learning-Based Forecasting on Ireland's West Coast**
Word Count:	6089
Page Count:	19

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:	Sabarinathan Chandramohan
Date:	16th September 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):	

Enhancing Wave Energy Conversion Efficiency through Deep Learning-Based Forecasting on Ireland's West Coast

Sabarinathan Chandramohan
x22213333

Abstract

This research aims to Enhancing the Wave Energy generation system efficiency to increase its power output and finding a new location on the west coast Ireland around Co. Sligo region to implementing the new wave energy generation plant to mitigate the energy crisis in the Republic of Ireland, in the western coast of Ireland due to the harsh North Atlantic weather creates a rough and high wave currents which is more suitable for the Wave energy generators to produce a continuous and steady electricity, this research study is divided into two phases in the phase 1 the data of test site plant of Co. Galway is collected and optimize the power output using the deep leaning-MLP, it's been observed from the test site of Co. Galway its maximum produced power is 5MW after optimization it's been predicted that more than 10MW energy can be produced and in the phase 2 this research aims to find a new location for laying this new optimized model, for this its uses the data from Buoy M4 and forecast the Wave height and Wave period by SARIMAX, by this the research achieve the its aims to reduce the energy crisis in Ireland.

1 Introduction

1.1 Research Background

In the growing global warming situations in the world leads to the over depletion of the ozone layer, many countries take an initiative to change more renewable source for producing the power, Europe is the largest consumer of the renewable energy resources like off-shore wind and solar farms but still from these sources the continuous and steady electricity can't produce, to overcome this need an Offshore Renewable Energy-ORE, these off shore farms are set up 10 to 15 kms (about 9.32 mi) from the coastal area as per O'Connell et al. (2024), this research aims to optimize the already presented wave power generation system and find the new location for optimize wave power generation system in the west coast of Ireland, this research uses the data from the test site plant OWC-Oscillating Water Column of Co. Galway and optimize it using the MLP model and extracting the data from the Buoy M4 of Co. Sligo and forecasting and predicating the future wave heigh using ARIMA/SARIMA model as the results a best suitable location and best power output are achieved by this research.

1.2 Research Question

How can the integration of deep learning method and precise wave period and heights forecasting can improve the operational efficiency of wave energy generators in the west coast of Ireland?

1.3 Proposed Solution

For addressing the proposed question this study split into two phases in the phase 1 it aims to improve the power generation on the existing wave power generation plant on west coast-test site of Co. Galway which is a OWC-Oscillating Water Column plant and by implementing the deep learning model utilizing the MLP the optimize power generation system can be achieved, and in the phase 2 which aims to find new location for the Wave energy generator to be laid, the region should have height waves like wave period and wave height should be more, for this best fit region on the west coast is Co. Sligo, the data from the M4 buoy collected and with that this research uses the SARIMAX and forecast the wave height and wave period after that predicted wave height can be used as the input for the optimized power generation model which acquired from the phase 1, by this it can answer the proposed question and prove its novelty.

1.4 Novelty of the Study

This research has a unique novelty, which directly states the propose and need of the more reliable power generation system in the Republic of Ireland, since on the west coast of Ireland there are not much wave energy generation plants are located, this research analysis the nature of west coast and its harsh waves pattern are best suitable for the continuous and steady power generation, this research is distinct in its head-to-head comparison and executed within the unique constraints.

1.5 Document Structure

The research is well structured and documented, starting with the introduction where this study explains the research question and motivation of this study, then followed up with the a detail study of related works, this research uses CRISP-DM and Scrum Baijens et al. (2020) for the implementation process, in the Implementation section it explains the data preprocessing, feature selections, model training and evaluation the train model with the test data, after the model have been evaluated with the evaluating matrix like r2, mape, mse to validate the accuracy of prediction of the research.

2 Literature Reviews

2.1 Wave energy power optimization through deep leaning

In the first phase of this research its aims to find out best optimize wave energy generation system using deep learning, for this a detail examination of the wave energy optimization through deep learning is viewed in this literature review, the study by Xiangyu and Zhisheng (2023) in 'Research on Power Forecasting Model of Wave Energy Generation Based on GRU Neural Network' the author create a wave power generation prediction model using neural network- GRU, this study focused in need of demand of accurate wave energy predictions and to resolved the significant fluctuations in wave energy, author used the GRU model since it has a greater impact in gradient issues compared to other RNNs, though this model performed well still its lacks in accuracy as more wave power and limited data for validation, the further accurate wave power predictions to optimize the WEC has done by Veurink et al. (2023) in 'Impact of Wave Forecasting Accuracy on Design and Power Predictions of a WEC Array' the author of this research creates a Wave forecasting model for the WEC-Wave Energy Converter arrays to improve the accuracy, this research also addresses the significant discrepancies in energy storage sizing and power delivery due to prediction errors, insufficient regional data and the unpredictability of wave currents are major limitations of this research, the increase in accuracy of the wave power generation with the different climate changes are addressed on the 'A data-driven long-term metocean data forecasting approach for the design of marine renewable energy systems' by Penalba et al. (2022), the author of this research develops a integration model of combining the machine learning and the forecast model for creating the long-term metocean data for marine renewable energy (MRE) systems, in this the author build three ML models RF, SVM and ANN to predict the power output, though this model perform well in predicting the wave energy generation power output it has few limitations like it fails in predicting the peak values, in the paper titled 'Deep Learning Prediction for Rotational Speed of Turbine in Oscillating Water Column-Type Wave Energy Converter' by Roh and Kim (2022), address the predicting the peak values and handle highly variable wave energy inputs by predicting the rotational speed of a turbine oscillating water column WEC to improved rated the power control, the author compares various performance of ML methods like MLP, RNN, LSTM, and CNN, out of which the LSTM performance well and provide more accurate predictions, it lacks in variability in prediction accuracy on less availability of the dataset, the study titled 'A predictive machine learning model for estimating wave energy based on wave conditions relevant to coastal regions' by Hassan et al. (2024) address the issue in the renewable energy demands in the coastal regions of Egypt, the author uses the machine learning model to predict the wave height and period for this XGBoost have been used, this model have achieved 0.52 rmse values, thought the predictions results are good the model still lacks on the historical data and performance variability, the research titled 'Multi-Step Long-Short Term Memory (LSTM) Time Series Ocean Waves Forecasting Model for Wave Energy Converters (WEC)' by Iqbal and Mehran (2023) address the need for accurate power predictions in the offshore Wave energy plants, the author uses LSTM to forecast the wave conditions for WEC, this study uses the real time data but still it lacks in use of larger dataset thought its handle sequential data effectively, improving power forecasting for WECs, in the paper 'Intelligence Visualization for Wave Energy Power Generation' by Liu et al. (2022) address the demand of the power and need of a platform monitors wave power indicators, the author creates a visualization platform for monitoring the power output and using the LSTM model for predicting the wave energy, the choice of LSTM is due to its ability to handle sequential data, the research have achieved in improving overall efficiency, but it did not address the problem of using the real sensor data for accurate modelling, by these review of various paper gives that MLP-Deep learning model is accurate in optimizing the wave power generation.

2.2 Predicting the future Wave height and patten

On the second phase of this research is to Forecast and predict the future Wave height and period is examined in this section of the literature review, the paper title 'Hybrid deep learning model for wave height prediction in Australia's wave energy region' by Ahmed et al. (2024) address the real-time wave height prediction, the author uses a hybrid deep learning model (CLSTM-BiGRU) to predict significant wave height, the models captured both temporal and spatial wave pattens, model perfomace stable in the wave prediction but lacks in lack of long-term prediction, the reliable wave height forecast and the long-term predictions are explained by Raj and Prakash (2024) in the research 'Assessment and prediction of significant wave height using hybrid CNN-BiLSTM deep learning model for sustainable wave energy in Australia' the author of this research uses the hybrid MVMD-CNN-BiLSTM models for accurate predictions of the wave height, the author uses the historical oceanic dataset, this hybrid model shows a larger accuracy, this model performance well in the accurate predictions using the MLP but still has few limitations like it can't able to use for the long term predictions, in the paper 'Improved Hierarchical Temporal Memory for Online Prediction of Ocean Time Series Data' by Qin et al. (2024) address the issue for enhance long-term prediction accuracy, for this the author uses an improved HTM model by incorporating GRU, this research demonstrates a better accuracy and stability in the wave predictions but has few drawbacks like the model needs further parameter tunning, in the paper titled 'Statistical machine learning models for prediction of China's maritime emergency patients in dynamic: ARIMA model, SARIMA model, and dynamic Bayesian network model' by Yang et al. (2024) deals with the improved emergency response planning and resource allocation, the author uses ARIMA and SARIMA models to forecast the maritime emergency medical cases, this research has limitations like it uses a small dataset and did it not used the environmental data, for capturing the long team forecasting and improving the accuracy of the wave forecast is explained in the paper 'Improving Linear Models for Univariate Long-term Time Series Forecasting' by Sellami et al. (2024), in this study the author uses an advance MLP models to improve the long team time series dataset and achieved superior performance, but it faces few limitations like difficulty in predicting change points, as per these review of the paper it have found the ARIMA/SARIMA model performance well in forecasting and predicting the future waves.

2.3 Energy crisis in Ireland

During this current situation the energy crisis is the major problem in the Ireland, due to the more population and certain economic sanctions in Europe, to overcome this situation this proposed research aims to build a new offshore wave energy generators in the west coast of Ireland, energy crisis problem and solution are discussed in the section of paper review, the research titled 'Forecasting and predictive analysis of source-wise power generation along with economic aspects for developed countries' by Hasan et al. (2024) discussed about the various aspects of future energy trends of developed countries, the author using a comprehensive study for forecasting the various different power generation source with its economic stands, in the paper '"We could have been leaders": The rise and fall of offshore wind energy on the political agenda in Ireland' by Roux et al. (2022) discuss about the political challenges and the developments for the off-shore farms in Ireland, this research speaks about the usage of the off-shore wind farms and its limitation in power supply and need for the wave energy generation system, the author gave clear picture of the energy surplus in Ireland and it can be mitigated through the more renewable energy sources.

2.4 Research Gap and Research Niche

The proposed research has a unique research question which aims to optimize the wave power generation and deploy a new wave power generation site in the west coast of Ireland in Co. Sligo using deep learning and forecasting of wave height, in the paper titled "An advanced geospatial assessment of the Levelized cost of energy (LCOE) for wave farms in Irish and western UK waters" by O'Connell et al. (2024) deals with the Levelized Cost of Energy (LCOE) for wave farms in Irish and western UK by addressing financial feasibility and optimal locations for the wave power generation plant, this study target to reduce uncertainty around wave energy projects, the author explains and finds that proximity to ports and the WEC grid positions are main objective of financial ability in the Ireland.

These literature reviews of various related research of Wave energy power optimization through deep leaning in 2.1 and predicting the future Wave height and patten in 2.2 sections, gives a detail idea of the current wave power generation and its power optimization through the deep learning techniques.

3 Methodology

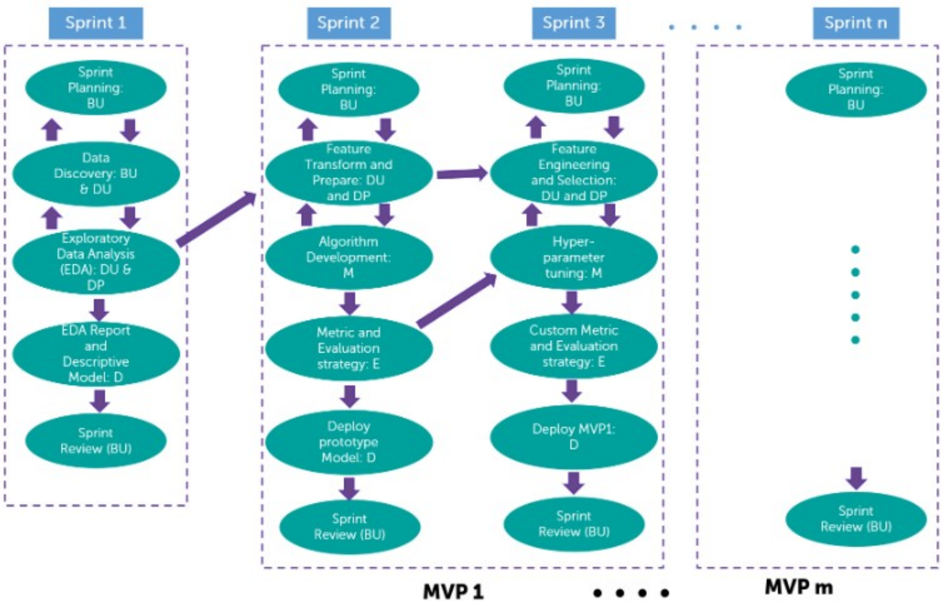

Figure 1: CRISP-DM & Scrum Methodology Source: Baijens et al. (2020)

The main objective of this research is to implement a new wave energy power generation plant in the west coast of Ireland by analysis in the western coastal areas which has more wave tides and easy port access proximity to reduce the cost for maintenance, this can be achieved by optimizing the power output of already present wave energy generation test plant in Co. Galway using the deep learning methods from the series of review of the related papers this research have concluded in using the MLP, and also this research have an objective to forecast the wave height and period in the coast areas of Co. Sligo which receives the hash tides of North Atlantic ocean, the cost of Co. Sligo have been chosen as the best location for the new wave energy generation plant to be laid with the new optimized system, for forecasting and predicting the future waves the SARIMAX model have been used in this research.

A new methodology have been used in this research is Iterative CRISP-DM (Figure 1), this methodology is an combination of Agile and CRISP-DM, as its name suggests this methodology is an Iterative approach as same as Agile model, where all the model and evaluation have been done in an Iterative approach, as this research follows CRISP-DM the cycle starts from Data collection, Data preprocessing, Data Preparation, Modelling, Evaluation & Prediction and Deployment & Results, this research uses ETL as for their Data Preprocessing.

3.1 Data Collection

The data gathering is the main and more crucial part as it needs to ethically obtained by following all the guidelines, since this research uses the public available data's it's come under the ethics standards, For this research the historical data's of both test site wave generation power plant and buoy data's are needed, the data's should be collected form the government sites or marine data's since its using the Realtime data from the buoy, the collection process should be legitimate with ethical considerations.

3.2 Data Understanding and Preprocessing

Once after the data's are collected it will be undergo with the serious of data understating as it very much needed before building the model, for better understanding the data a detail EDA- Exploratory Data Analysis have been taken with detail visualization of various variables have been done, ETL have been utilized in this project since its using the real time data, in ETL under the Transform part the data are Transformed like cleaning, merging for more understandable have been done and then it's Loaded into PostgreSQL as the Data Warehousing, for this ETL-Dagster tool have been used and last the data's are pre-processed and moved it in PostgreSQL as Data warehousing.

3.3 Data Preparation

After the data are understood and pre-processed and stored in the PostgreSQL via ETL, the data must be prepared for modelling, in the Data Preparation the target variable are found by Feature extraction and Feature selection have done as part of Data preprocessing to reduce dimensionality and improve computational efficiency.

3.4 Modelling

This research has divided into two phases, In the phase 1 the optimizing of the power output of the wave power generation system is done using MLP and in Phase 2 the forecasting of the wave height and the period of the Sligo coast is predicted by using SARIMAX, for the 1st phase the feature is selected and the target variable is defined and model is fit, the MLP model is fit with grid search with cross-validation and various Hyperparameter like early stopping and random search, for the phase 2 the feature is selected with the target variable the ARIMA model is fit with the exogenous model to improve the predicting accuracy the hyper tunning is done for improved accuracy.

3.5 Model Evaluations & Predictions

The models are evaluated based on the training and validation loss, which was not done in the model training part, in the Evaluation of the fit model gives the better understanding of the model performance, few statical evaluation metric are determined in order to find the performance accuracy of the model, the evaluating metric like R2, RMSE, MAPE and MSE are determined to find the accuracy this was motivated by the research by Klemm et al. (2023), if the accuracy is not to the mark further hyper tunning is done to increase the accuracy and sets a benchmark, once the model is evaluated with the test data the predictions are done, with that predictions it ensures the reproducibility and which is directly aligns with the research aim, those predictions are drawn in a graph for better understanding, These metrics provided a clear view of the models' performance, considering both the positive and negative.

3.6 Deployment & Results

As for the final step all the results which gathered are consolidated and the model performance and predictions are made the analysis, and the result outputs are the evidence for the proposed research objective which is to implemented a new wave generation power plant in the west coast of Ireland by forecasting and predicting the future wave pattens.

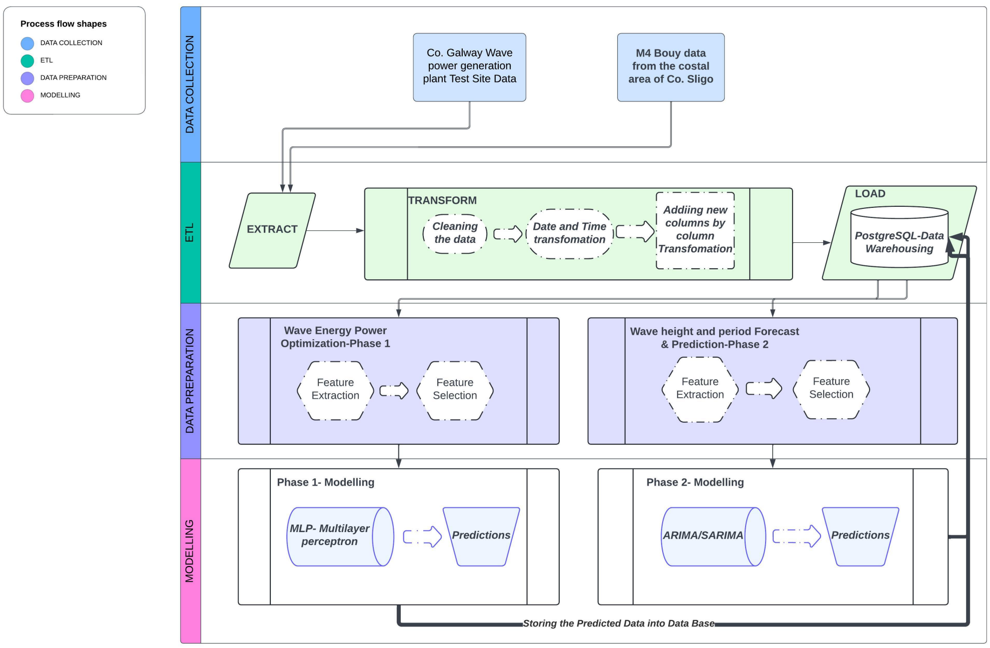

Figure 2: Research Architecture

4 Design Specification

This research aims to optimize the power output in the wave energy generation system and implement a new wave energy generation system in the west coast of Ireland by advance forecasting of wave height and period in the coastal region of Co. Sligo, since its objective of this research uses the real time data's which is collected from the test site wave energy plant from Co. Galway and wave patten data from M4 buoy, this research uses ETL-Extract, Transform and Load and storing in PostgreSQL database for data warehousing, the research design architecture(Figure 2) includes, data collecting where the raw data's are collected and then ETL for the Data preprocessing, for ETL this research uses Dagster tool, where the Extract, Transform and Load have implemented, in the Transform stage the data's are transformed and cleaned and then loaded into PostgreSQL as the Data warehousing, after that Data preparation is done for the model building, as per the research objective the model is built for two phase for phase 1- Optimizing the wave energy power generation using MLP is built then the forecasting and predicting the wave pattens using SARIMAX is done as part of phase 2 of this research, after that those predicted values data are stored back into the PostgreSQL for the future.

4.1 Techniques and Architecture

4.1.1 ETL Architecture

Dagster is used as the ETL tool in this research, a detail ETL pipeline is define for this research to handle the real time data's and to reduce the time for the data preparation time this will allow an organized, automated and efficient way in the data preparation, this automated system will provide reproducibility in the future implementations with efficiently, in detail Dagster ETL structure is shown in the Figure 3, the separate process

Figure 3: ETL Dagster Architecture

of Extraction, Transform is defined where the raw data is collected and transformed then its joined using concatenation and Loaded into PostgreSQL for the Data warehousing, after the prepared data is stored in the PostgreSQL its then used for the analysis and modelling.

4.1.2 MLP-Multilayer perceptron Architecture

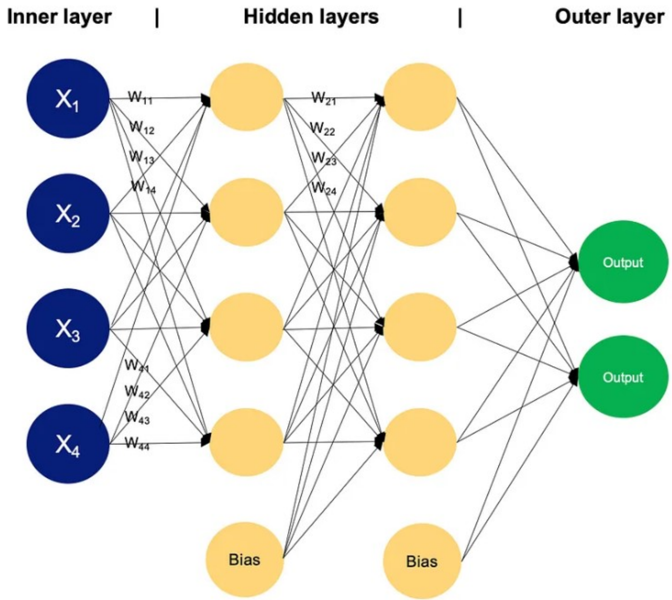

Figure 4: MLP Architecture Source:M'zoughi et al. (2024)

Optimizing the efficiency of wave power generation system is done on the 1st phase of this research, for this the MLP- Multilayer perceptron have been implemented, the choice of using MLP is as its part of ANN which has high potential in handling for the regression problems, this proposed research is all about finding the factor which is directly related with the power output and improving them to increase the power efficiency, for this regression the MLP is used, the MLP has three layer in its neural network, an input layer, hidden layer and an output layer(Figure 4), due to its high computationally powerful the MLP model perform well in these conditions and provide better accuracy.

4.1.3 SARIMAX Architecture

For forecasting and predicting the Wave pattens like wave heights and periods for the phase 2 of this research, SARIMAX is used for forecasting Seasonal Auto-Regressive Integrated Moving Average with Exogenous variables, which is advance model of native ARIMA model which using an extra variable for the more efficient forecasting which was used by Rachmanda and Adytia (2023) and this motivates this research to take SARIMAX, this is research the SARIMAX model is trained and fit with the train data and forecast the out-of-sample estimations, with that estimation the in-sample estimate predictions are made and the with that the future wave heights can be determined as the part of the forecast, and this model of predicting using SARIMAX is more efficient and accurate in the forecast predictions, since research has many independent variable are dependent on the target variable its necessary for Multi-variable forecasts, for this multi variables forecast can be done precisely using SARIMAX where the variables are transformed and used as the exogenous variables and the forecasting is done and which this model performance well in the predictions.

4.2 Implementation Requirements

4.2.1 Hardware Specifications

Need an higher performing computer or system with an operating a multi core and higher efficiency CPUs for handling an larger datasets for training and testing of those models, since this research uses the real time and historic data's it need an multi core and more than 16GB RAM, also for the less time for the training models GPUs are needed, the GPU with NVIDA CUDA are much required for the MLP model to reduce the training time by parallel processing.

4.2.2 Dependent Libraries and Tools Used

Python is the programming language has been used in this research, since it exabits a good ecosystem for machine learning libraries, for the implementation this research uses all the desired machine leaning and deep learning libraries and frameworks like pandas, Dagster, TensorFlow, scikit-learn, kerars and visualization libraries and various tools have used as part of this research those are described in the below table.

Tool	Purpose
Jupyter Notebooks	Interactive coding and building the models and testing the algorithms
PostgreSQL	Data Ware housing and used the database for storing collecting the data
Dagster ETL tool	Data Preparation and for pipeline management

5 Implementation

5.1 Data collection

The first step in a machine leaning or deep learning implementation is data acquiring and collecting, for this research the data are collected and acquired from all the government and public available sites, the wave energy generation dataset from the test site from the Co. Galway, wave energy power generation is collected from the Sustainable Energy Authority of Ireland - SEAI from the ocean energy Ireland and data.gov.ie websites, from these sites the real time data of the wave energy power plant is acquired which can be used for optimizations for the wave energy power generation efficiency, for the wave patten dataset its collected from the marine.ie and met.ie government site, the data from M4 buoys is collected in Sligo coast, from these this research utilized those data for the optimization power generation and wave height forecast, along with that collection of the dataset this research also focuses on the ethical compliance, following all the ethical guidelines via collection and using the data this research follows all the principles of the ethical guidelines, Legitimate data uses and governance standards are followed in this research the dataset serves as a tool for developing and testing for the wave power generation optimization system, also proved a practical learning for overcome the energy crisis of the Republic of Ireland.

5.2 Data Preparation

After the data are collected it must be analyzed and prepared for the modelling, for this the research uses ETL for the data preparation, the choice of using ETL for this research is usage of the real time data to handle this an automated system is needed for the data preparation, Dagster tool is used for the ETL process, this reseach analysis the ETL work done by Alak et al. (2023) and motivated and constructed the ETL and by which created a detail pipeline for the ETL (Figure 5) where the data are Extracted, Transformed and Loaded into a database for this PostgreSQL is used as the data ware housing and the database center for storing and handling the data, in the Extract phase

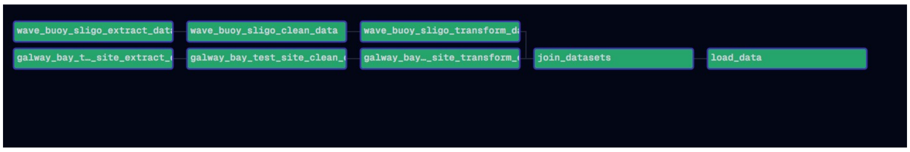

Figure 5: ETL Pipeline Visulization

the data from these cite are downloaded as the CSV file, in the Transform phase the raw data are under goes various transformation to make the dataset are structure and easy for analysis, for this date time transformation, converting data types, normalizing data, creating new calculated columns and data cleaning in which it checks the handle missing values and ensure that numeric columns are correctly formatted and also certain new features are created like wave_height and Wave Period for the better analysis and this will provide reproducibility in the future implementations with efficiently, after that the transformer data is Loaded into the PostgreSQL for the analysis and model building.

5.3 EDA-Exploratory Data Analysis

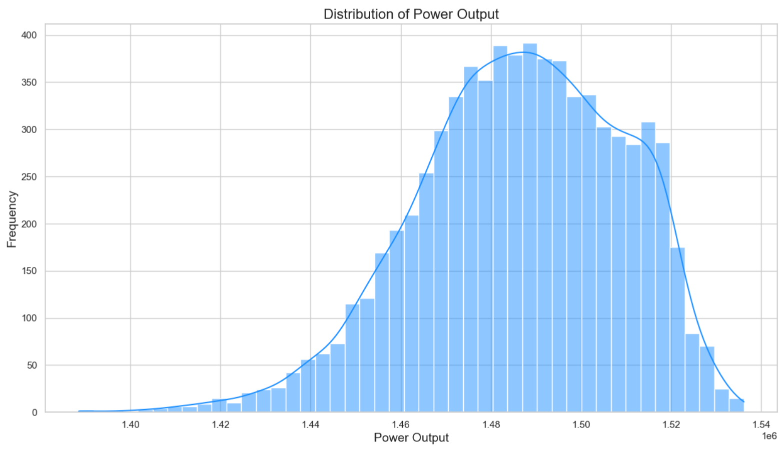

Figure 6: Wave height vs power output

In the Exploratory Data Analysis step a detail over view or an understanding of the data is examined, from the series of statical analysis of the data and from the various visualization the exact information on the dataset is acquired, this research main objective is to optimize the wave power generation efficiency through the forecasting the wave patten in the Co. Sligo area to eliminate the energy crisis in the Ireland, for this it have proposed a two phase implementation in the 1st phase the optimization is done using MLP and in 2nd phase forecasting is done using SARIMAX, for these this two separated implementation is used in this research.

5.3.1 Phase 1 Optimization of Wave Power Generation Efficiency

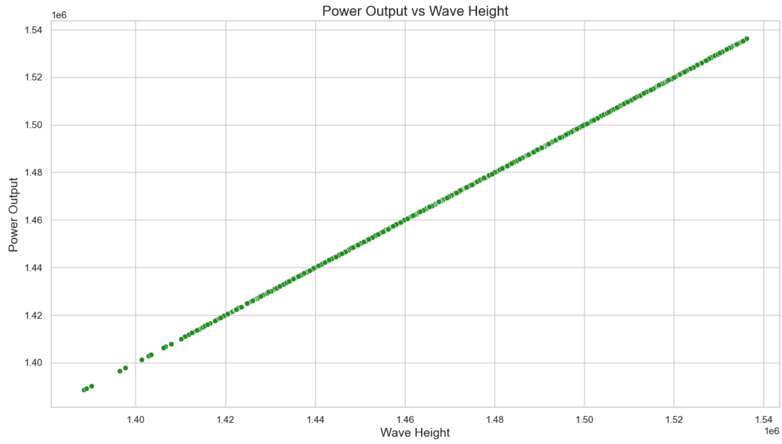

Figure 7: Scatter plot of power output vs wave height

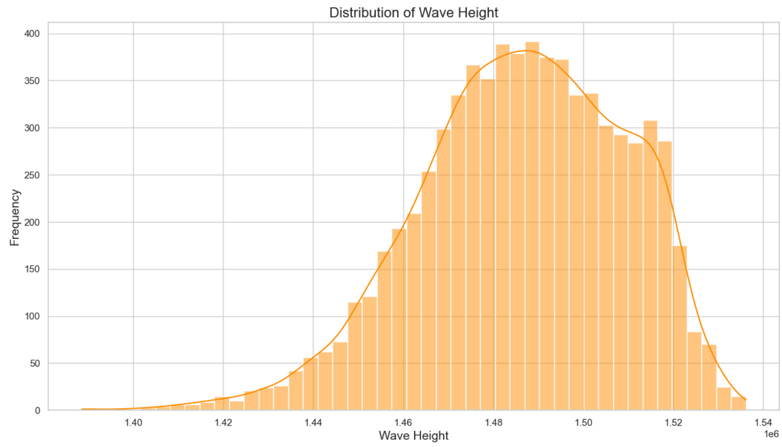

Figure 8: Distribution of Wave Height

In the optimization of the wave power generation system these columns are used and used, a SELECT query these columns are retrieved from the PostgreSQL database, the columns are '"X1", "X2", "X3", "X4", "X5", "X6", "X7", "X8", "X9", "X10", "X11", "X12", "X13", "X14", "X15", "X16", "Y1", "Y2", "Y3", "Y4", "Y5", "Y6", "Y7", "Y8", "Y9", "Y10", "Y11", "Y12", "Y13", "Y14", "Y15", "Y16","Power_Output", "wave_height' in which the 'Power_Output' is the target variable the basic knowledge or understanding of these variables are X1, X2,X3..etc are the power generation output of right WEC and Y1,Y2,Y3..etc are the power generation output of left WEC of the test site wave power generation plant, at the EDA the deep understating of the dataset and the selected variables are analyzed like brief structure and description are analyzed with the statical analysis of understating the mean, median and mode, then visualization are drawn for the better understanding of relation with the target variable with other variables, in the Figure 6 a histogram is drawn for 'wave height vs power output' from this it's clearly understood that greater wave height produce high power and the Figure 7 the scatter plot also explains the same increase in power due to the wave height and in Figure 8 the histogram plot explains the variations of 'Distribution of Wave Height'.

5.3.2 Phase 2 Forecasting the wave height and period

For the Wave height and period forecast, these columns or variables are used, by using the SELECT query these variables are retrieved from the PostgreSQL database, for forecasting the variables from the M4 buoy dataset are utilized after Transformed, the columns such as '"wave_height", "PeakDirection", "SignificantWaveHeight", "Hmax", "SeaTemperature", "MeanCurSpeed", "MeanCurDirTo", "WavePeriod", "time"' are used and in this 'SignificantWaveHeight' is the target variable used for forecasting the wave patten and predicting the future wave height, the variables are much easier to understand and it collected from the real time M4 buoy which is located in the coast of Co. Sligo, the

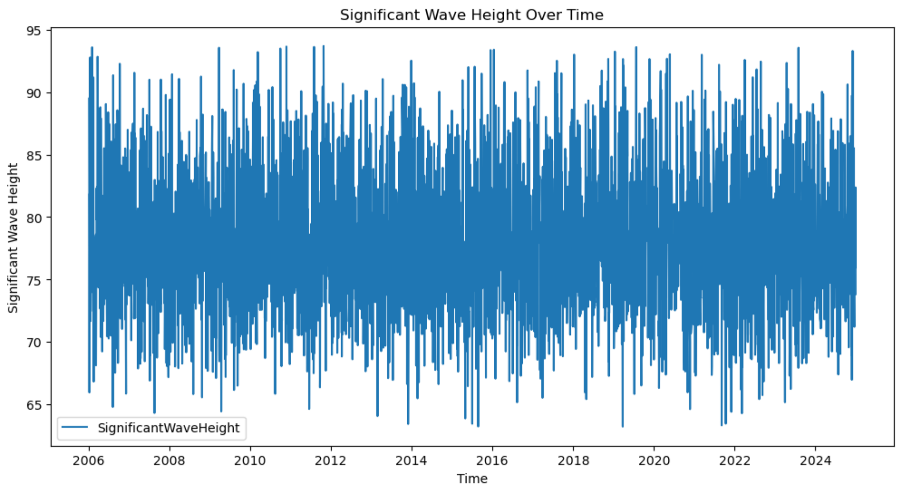

Figure 9: Significant Wave Height

EDA is performed to analysis the time series structure of this data, converting the time into date stamp, handling missing values and remove the outliers are done as part of the cleaning and detail statical analysis like mean, median and mode of this dataset are described for the greater understanding of its structure, a line plot is drawn in the Figure 9 to represent the detail timeline of the wave height from 2006 till 2024, from which it can be observed that how the wave height is differ over the time period.

5.4 Feature Extraction and Feature Selection

5.4.1 Phase 1 Optimization of Wave Power Generation Efficiency

For the optimization after the EDA is done now this research takes the feature extraction and selection for the model building, for this all the necessary variables are extracted from the data base and the target variable 'Power_Output' is set, for accurate feature extraction the data must be cleaned as part of the feature extraction, then in the feature selection is done by correlation analysis, for this correlated matrix is created (Figure 10) to find the relationship between the variables and to identify the variables which is strongly correlated with the target variable, as part of the feature selection the selected feature

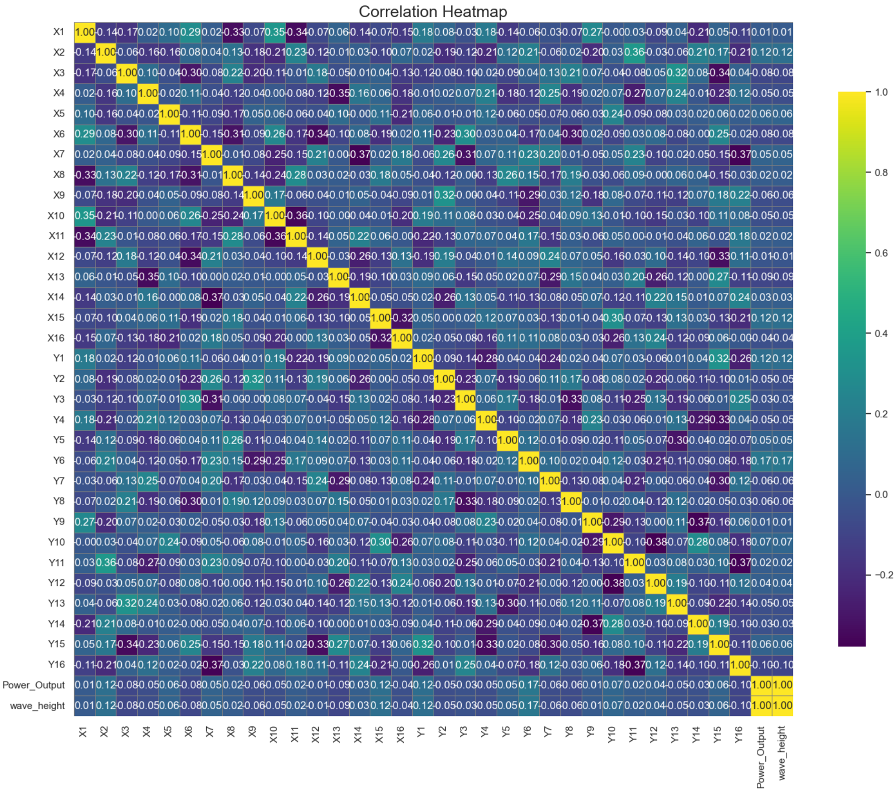

Figure 10: Correlation Heatmap

variable or highly correlated variable are combined to reduce dimensionality using the dimensionality reduction technique, after the feature is selected this research uses feature scaling using 'MinMaxScaler' to normalize the data for making all the feature are contributing equally in the model training.

5.4.2 Phase 2 Forecasting the wave height and period

The data are extracted from the database for the forecasting of the wave pattens, the relevant features are retrieved from the databased and 'SignificantWaveHeight' is set as the target variable, the missing values in the dataset are handled using the interpolation for maintaining the continuity in the time series then with the target feature the outlier are removed using the Interquartile Range (IQR) by which the data quality can be improved, after the feature is extracted for the selection feature in the time series data the feature engineering have been used in this research, by using moving averages 'SWH_MA7' and 'SWH_MA30' are created to capture the short-term and long-term trends by this it helps in the prediction accuracy, also the seasonal pattens are extracted based on the other parameters and using the seasonal variations in the data for better forecasting.

5.5 Model Development

5.5.1 Phase 1 Optimization of Wave Power Generation Efficiency-MLP

Optimization the wave power generation system of the test site wave energy generation plant of Co. Galway need the MLP as part of this research proposed, a neural model is created using the Keras Sequential API for the MLP, which has three layers have been created including the Dense, Dropout and Activation functions, this MLP model has 64 layers and 32 neurons to handle the large data and to provide the accurate results and created a single output neuron for the regression to optimize the power output, this research also created the hyperparameter tunning to increase the performance accuracy, using the 'Random Search' of the Keras Tuner's the hyperparameter tunning is done, the use of random search is to search the best parameters of the hyper tunning by this the model accuracy are optimized, the MLP model is compiled using the 'Adam' optimizer and MSE for the loss function, since this is an regression tasks the MSE remove the large error significantly, then the model is trained with the train data for this the data is split into 80-20 split for the model training, 80% is consider as the train data with that the model is trained and to prevent the overfitting early stopping and learning rate reduction callbacks are used in this research, after the model is trained with the test data the remaining 20% data is test data which is used to validated and check for the model performance, the metrics like MAE, MSE and R2 are used to provide the accuracy and performance of the model.

5.5.2 Phase 2 Forecasting the wave height and period- SARIMAX

For the forecasting and predicting the wave height in the coast of Co. Sligo this research uses the SARIMAX as the model for this better handling of the time series forecast, since the data has both regression and seasonal feature SARIMAX is best in capturing the both, to find out the best combination of parameters the grid search is used in this research of the range p, d, q by using the ARIMA and seasonal components, by this the validation error can be minimized, then the SARIMAX model is fit with the training data with the exogenous variables which has the seasonal feature to improve the predictions, once after the mode is fit checked for its p-value, the variable whose having p-value greater than 0.05 are removed and again the model is fit and checked again for p-value now this model is consider as the final trained model, then the forecasting is done with the fit model, after that the model is validated with the forecasted values, the actual values are compared with the forecasted values and also with the forecast accuracy over a specific time period for the future predictions.

5.6 Result Documentation and Finalization

The results and outcomes of both model evaluation are compiled and analyzed as for the finalizations of the research, all the evaluation scores like performance and statistical analysis, for the wave power generation optimization the metrics for the performance validation are MAE, MSE and R2 scores are calculated and for Wave height forecast in the Co. Sligo coast the MSE and R2 are calculated and also the forecast accuracy is visually represented with comparing actual vs predicted values.

6 Evaluations

6.1 Overview of Results

The objectives of this research have been achieved, optimizing wave power generation efficiency and forecasting wave height and period, MLP neural network was created to

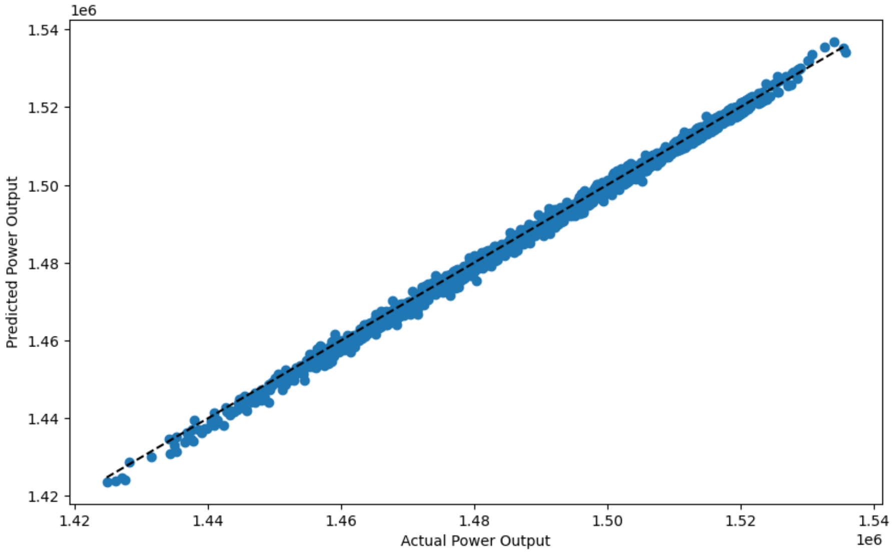

Figure 11: Actual vs predicted Power_Output

optimize the wave power generation output using wave height and with SARIMAX forecast of wave height using seasonal trend was implemented, for the phase 1 evaluation series of visualization is drawn, a scatter plot of actual vs predicted Power_Output are plotted is Figure 11 explains the model is well aligns with the actual data, for the phase 2 few plots are made for the model evaluations, the forecasted values are drawn with actual values are is plotted in Figure 12, both model performance well and exhibits 97% accuracy in optimizing and predicting.

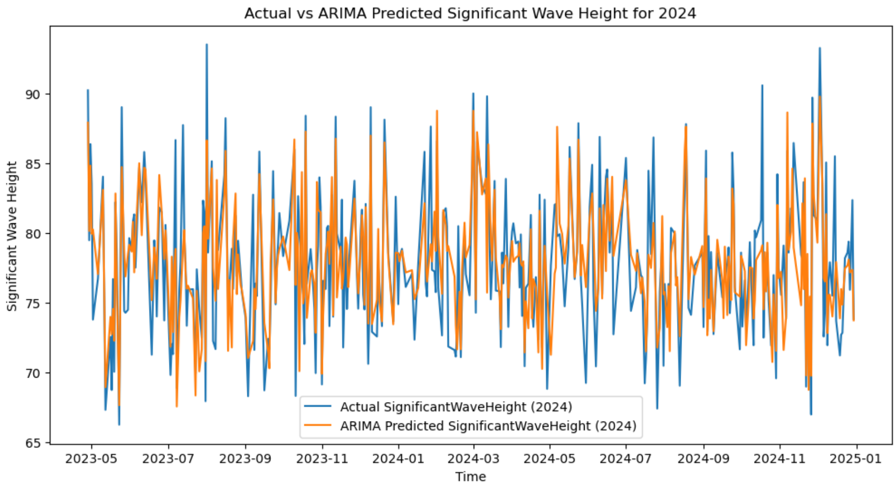

Figure 12: Actual vs SRIMAX Predicted Significant Wave Height

6.2 Statistical Evaluations

The MLP-neural network model for optimizing wave power generation system achieved an MAE of 2.3, MSE of 5.7 and R2 score of 0.92 a detail comparison is given in the table 1, Error distribution KDE histogram plot (Figure 13) shows a minimal bias and variance, which indicating the effectiveness in optimizing the wave power output.

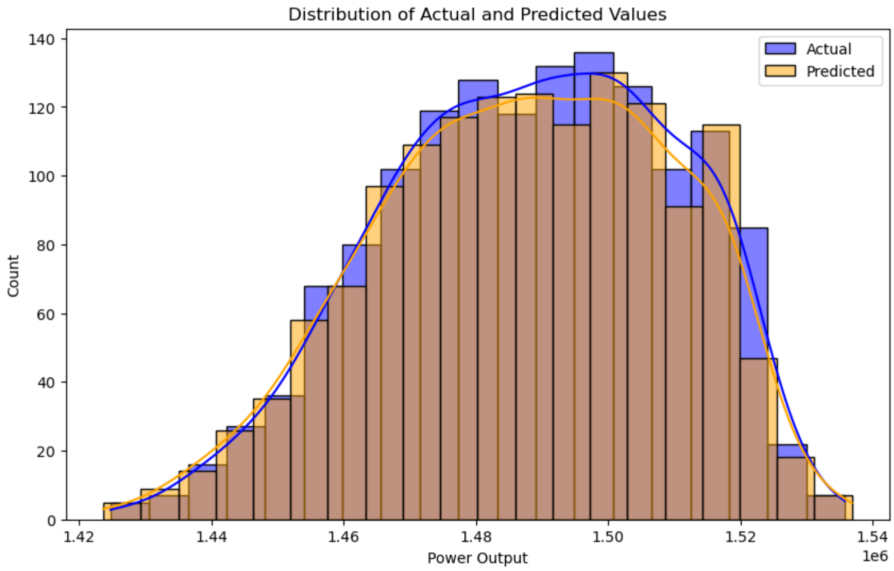

Figure 13: Distribution of Actual and Predicted power ouput

Table 1: Phase 1-Optimization of Wave power energy generation

Evaluation Metrics	Values
MAE-Mean Absolute Error	2.3
MSE-Mean Square Error	5.7
R-squared	0.92

The wave height and patten forecast using SARIMAX model was well fit, and predictions are made accurately this was validated from the actual vs predicted significant wave height forecast plot (Figure 14) and achieved MAE of 0.15 and MSE of 0.35 scores, detail comparison is given in the table 2

Table 2: Phase 2-Wave height and period Forecast

Evaluation Metrics	Values
MAE-Mean Absolute Error	0.15
MSE-Mean Square Error	0.35

6.3 Discussions

The MLP neural network model build for optimization of wave power generation system performed well effectively increase the power output of the wave power generation plant, its accuracy in optimization and with the evaluation metrics its understood that model is well performed, and the SARIMAX for the wave height forecast and predicting the future wave pattens in the Co. Sligo was predicted well and providing reliable predictions with better accuracy, these models are build and producing promising results as per the research proposed have significant for practical applications in wave power energy optimization and implementation of new wave plant in Co. Sligo, which will directly achieve the objectives of the research.

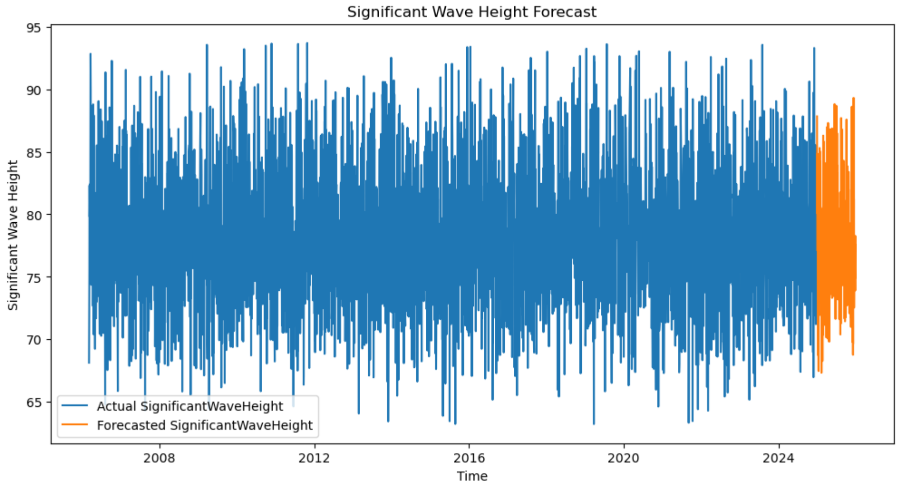

Figure 14: Significant Wave Height Forecast

7 Conclusions and Future Works

As this research concludes successfully in achieving its proposed solution, by developing the fully efficient models using deep learning, optimizing wave power generation efficiency and forecasting wave height, the evaluation results provides a greater insight of the model performance, to the future enhancement involves as this research creates a new milestone in the off-shore wave energy farm for the Republic of Ireland from this the next steps can be easily achieved, a practical or physical wave energy plant can be implemented in Co. Sligo based on this research output and to the future works a simulated study work can be done to evaluate improved power output, which can be a seed for designing a new power plant in Co. Sligo to reduce the energy crisis in Ireland.

References

Ahmed, A. A. M., Jui, S. J., AL-Musaylh, M. S., Raj, N., Saha, R., Deo, R. C. and Saha, S. K. (2024). Hybrid deep learning model for wave height prediction in australia's wave energy region, *Applied Soft Computing* **150**: 111003.

Alak, B., Revanbahş, A., Argün, N. and Çalım, A. (2023). Anomaly detection for etl packages runtime: A machine learning approach, pp. 108–113.

Baijens, J., Helms, R. and Iren, D. (2020). Applying scrum in data science projects, **1**: 30–38.

Hasan, S., Hossain, I. U., Hasan, N., Sakib, I. B., Hasan, A. and Amin, T. U. (2024). Forecasting and predictive analysis of source-wise power generation along with economic aspects for developed countries, *Energy Conversion and Management: X* **22**: 100558.

Hassan, M. K., Youssef, H., Gaber, I. M., Shehata, A. S., Khairy, Y. and A.El-Bary, A. (2024). A predictive machine learning model for estimating wave energy based on wave conditions relevant to coastal regions, *Results in Engineering* **21**: 101734.

Iqbal, S. and Mehran, K. (2023). Multi-step long-short term memory (lstm) time series ocean waves forecasting model for wave energy converters (wec), pp. 405–412.

Klemm, J., Gabriel, A. and Torres, F. S. (2023). Predicting future wave heights by using long short-term memory, pp. 1–10.

Liu, X., Liu, Y., Guo, J., Lou, R. and Lv, Z. (2022). Intelligence visualization for wave energy power generation, pp. 986–987.

M'zoughi, F., Lekube, J., Garrido, A. J., De La Sen, M. and Garrido, I. (2024). Machine learning-based diagnosis in wave power plants for cost reduction using real measured experimental data: Mutriku wave power plant, *Ocean Engineering* **293**: 116619.

O'Connell, R., Kamidelivand, M., Furlong, R., Guerrini, M., Cullinane, M. and Murphy, J. (2024). An advanced geospatial assessment of the levelised cost of energy (lcoe) for wave farms in irish and western uk waters, *Renewable Energy* **221**: 119864.

Penalba, M., Aizpurua, J. I., Martinez-Perurena, A. and Iglesias, G. (2022). A data-driven long-term metocean data forecasting approach for the design of marine renewable energy systems, *Renewable and Sustainable Energy Reviews* **167**: 112751.

Qin, T., Chen, R., Qin, R. and Yu, Y. (2024). Improved hierarchical temporal memory for online prediction of ocean time series data, *Journal of Marine Science and Engineering* **12**(4).

Rachmanda, A. A. and Adytia, D. (2023). Time series wave forecasting with gru with attention approach, case study in jakarta bay, indonesia, pp. 49–54.

Raj, N. and Prakash, R. (2024). Assessment and prediction of significant wave height using hybrid cnn-bilstm deep learning model for sustainable wave energy in australia, *Sustainable Horizons* **11**: 100098.

Roh, C. and Kim, K.-H. (2022). Deep learning prediction for rotational speed of turbine in oscillating water column-type wave energy converter, *Energies* **15**(2).

Roux, J.-P., Fitch-Roy, O., Devine-Wright, P. and Ellis, G. (2022). "we could have been leaders": The rise and fall of offshore wind energy on the political agenda in ireland, *Energy Research Social Science* **92**: 102762.

Sellami, M. A., Seyf Elislam Hamzi, O. and Khaldi, B. (2024). Improving linear models for univariate long-term time series forecasting, pp. 1–8.

Veurink, M. G., Weaver, W. W., Robinett, R. D., Wilson, D. G., Matthews, R. C., Michelen, C. and Dallman, A. R. (2023). Impact of wave forecasting accuracy on design and power predictions of a wec array, pp. 399–404.

Xiangyu, T. and Zhisheng, Z. (2023). Research on power forecasting model of wave energy generation based on gru neural network, pp. 606–611.

Yang, P., Cheng, P., Zhang, N., Luo, D., Xu, B. and Zhang, H. (2024). Statistical machine learning models for prediction of china's maritime emergency patients in dynamic: Arima model, sarima model, and dynamic bayesian network model, *Frontiers in Public Health* **12**.