First project:

You will have to form a group of 3 to 4 students.

Choose a dataset from the UC Irvine Machine Learning Repository (https://archive.ics.uci.edu/ml/index.php) with at least 5000 instances and 20 attributes for classification or regression. Compare how the different approaches seen in class perform on this dataset to predict accurately the classes or the values of the unlabeled data. You should determine what are the best hyper-parameters for each approach you are using. You could use any Python libraries.

<https://archive.ics.uci.edu/dataset/494/wave+energy+converters>

**Wave Energy Converters**

This data set consists of positions and absorbed power outputs of wave energy converters (WECs) in four real wave scenarios from the southern coast of Australia.

A close-up of a computer code

Description automatically generated

**Attribute Information**

Additional Information

Attribute: Attribute Range

1. WECs position {X1, X2, â€¦, X16; Y1, Y2,â€¦, Y16} continuous from 0 to 566 (m).

2. WECs absorbed power: {P1, P2, â€¦, P16}

3. Total power output of the farm: Powerall

**Size-Constrained Environment add in:** These 16 WECs are optimized within a size-constrained environment. This means that there are limitations on the physical size or placement of the WECs.

papers:

<https://arxiv.org/ftp/arxiv/papers/2011/2011.13130.pdf>

<https://www.youtube.com/watch?v=4yHHF9goNqE>

Wave Farm

<https://www.youtube.com/watch?v=-XWbVMtNnaw>

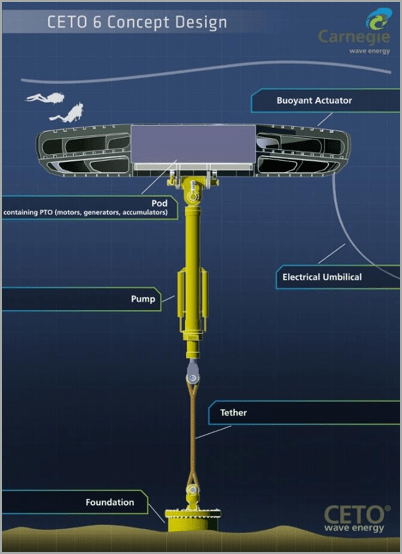
A group of yellow buoys floating in the water

Description automatically generated

A computer screen with graphics

Description automatically generated

CETO



Pages: 5 - 8 pages

- include a presentation of the research questions

- the chosen methods to tackle them

- a presentation of the results and discussion and a conclusion/future work

- attach a description of the participation of each student to the project

**First Project: Machine Learning Analysis**

**Abstract**

Delving into the innovative world of renewable energy and sustainable technology. This research project explores Wave Energy Converters (WECs) and their integration with machine learning analysis. The generation of power within a wave farm depends on the arrangement of WECs and the prevailing wave conditions. Crafting an effective WEC configuration is pivotal for maximizing power production. This document outlines the objectives, methodologies, and deliverables for the project, focusing on machine learning analysis in the context of WECs. In this context, we leverage data collected from real-world test sites to develop a model capable of predicting the wave farm power output. The project aims to explore and compare various machine learning approaches to predict and optimize the total power output of a farm of WECs in a size-constrained environment. This will involve utilizing machine learning algorithms to predict total power outputs based on the positions and absorbed power of WECs.

1. **Introduction**

In the present era, one of the most pressing global challenges revolves around the generation of energy while mitigating adverse environmental effects. Conventional methods of energy generation are reliant on fossil fuels and have proven unsustainable. Historically, renewable energy resources carried a higher cost burden in comparison to conventional fossil fuel sources. This led to deterring governmental interest in their promotion. Fortunately, this paradigm has recently undergone a noteworthy shift. Wave energy is a promising aspect of renewable energy, boasting substantial energy potential while causing minimal environmental repercussions. The efficiency of wave energy extraction is contingent not only upon the design of wave energy converters, but also on the strategic location of the wave farms. It revolves around the optimization of WECs within a size-constrained environment to harness wave power and convert it into electricity.

The project will employ machine learning algorithms to predict the total power outputs based on the positions and absorbed power of WECs,

1. **Dataset**

The dataset used in this study is sourced from the UC Irvine Machine Learning Repository and consists of data from four wave scenarios observed along the southern coast of Australia, incorporating locations such as Sydney, Adelaide, Perth, and Tasmania. The four strategic WEC locations have been methodically placed within the confines of a size-constrained environment. The dataset employed for analysis comprises a total of 48 attributes. Notably, the first 32 attributes pertain to the precise locations (both latitude and longitude) of the 16 WECs. The longitude and latitude are represented as X1, …, X16 and Y1, …, Y16 respectively. This is denoted as continuous values ranging from 0 to 566 meters. {add fig DistinctConfigurationOfTheWaveFarm } Next, the subsequent 16 attributes represent the power absorbed by these WECs. The final attribute summarizes the total power output generated by the entire array of converters within the wave farm. The positional data of the WECs are employed as essential inputs for the predictive modeling of total power output. This extensive dataset encompasses a total of 72,000 rows, each corresponding to a distinct configuration of the wave farm where the positions of the 16 WECs are varied. Their respective absorbed power levels are recorded for analysis and modeling.

[?] - https://arxiv.org/ftp/arxiv/papers/2011/2011.13130.pdf

1. **Data Preprocessing**

- Data preprocessing tasks will include handling missing values, encoding categorical variables, and scaling data for analysis.

1. **Model Selection**

- Multiple machine learning algorithms suitable for regression will be selected.

1. **Model Training and Evaluation**

- The selected models will be trained on the dataset.

- Evaluation of each model's performance will be conducted using appropriate metrics, such as accuracy, F1-score, and mean squared error.

1. **Hyper-Parameter Tuning**

- Identify the best hyper-parameters for each model to optimize their performance.

1. **Comparison of Results**

- Results and performance metrics of different models will be compared.

- The model(s) achieving the best results will be determined.

1. **Conclusion and Findings**

- A summary of findings from the analysis will be provided.

- Implications of the results will be discussed.

1. **Description of Student Participation**

-Attach a description of the participation of each student in the project.

**References**

[UC Irvine Machine Learning Repository]

Neshat,Mehdi, Wagner,Markus, and Alexander,Bradley. (2019). Wave Energy Converters. UCI Machine Learning Repository. <https://doi.org/10.24432/C5831S>.