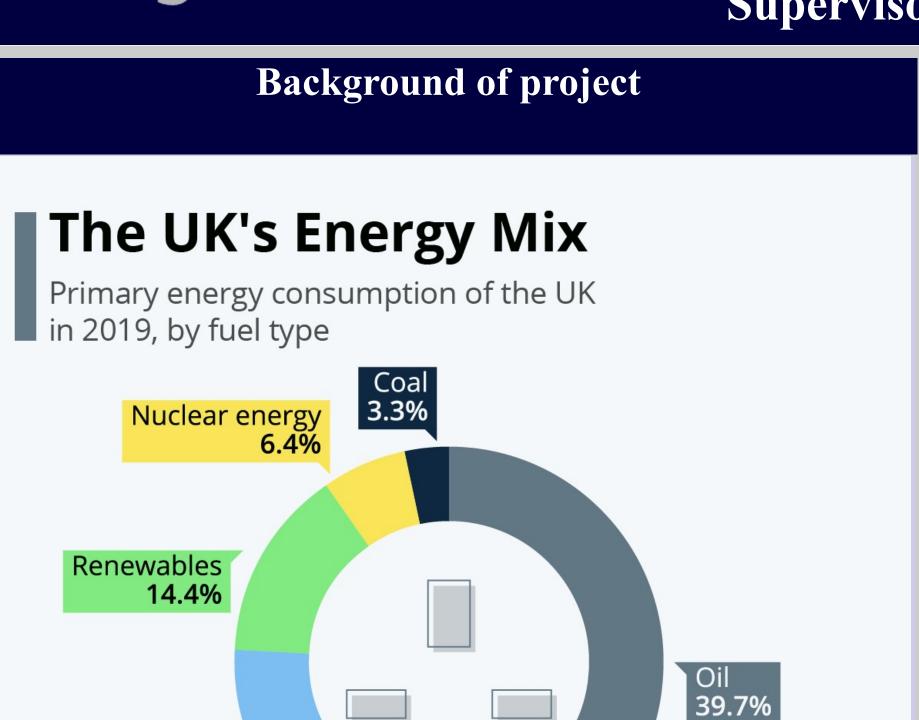


# Intelligent Data Analytics for Energy Producing and Energy Consuming Systems

by Khivishta Boodhoo







Primary energy comprises commercially traded fuels, including modern renewables used to generate electricity. Source: BP

Natural Gas

36.2%



Figure 1. UK total primary energy consumption in 2019.(Source:BP)

- . UK energy mix still comprises of a significant amount of oil and gas. Therefore it makes sense to optimise the energy produced/ used in this area
- . There is an upcoming growth in renewable energy particularly wind turbines and therefore optimizing energy produced is very

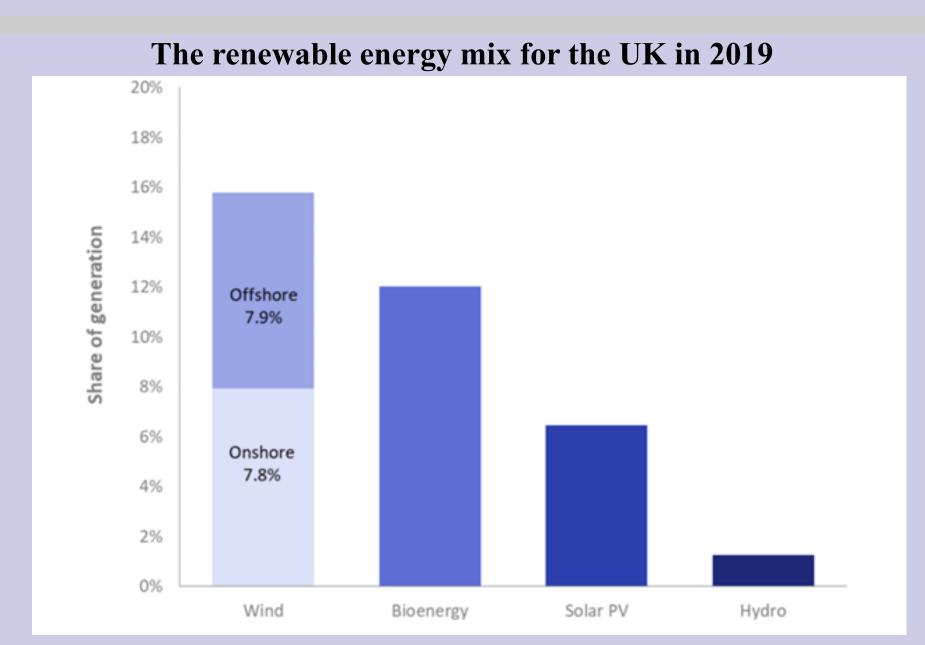


Figure 2.GOV.UK. 2019. Energy Trends: September 2019 - GOV.UK . [ONLINE] Available at https://www.gov.uk/government/statistics/energy-trends-september-2019. [Accessed 31 Mar



Figure 3. Offshore ORE catapult wind turbine in FIFE (Source: https://ore.catapult.org.uk/) [Accessed :31-Mar-2022]

- Optimise the production of energy from the offshore wind turbine.
- . More efficient operation.

Food Water Waste Research Group

• Project in line with UK goal to decarbonize the UK by 2050.

## **Expected Outcome of this study**

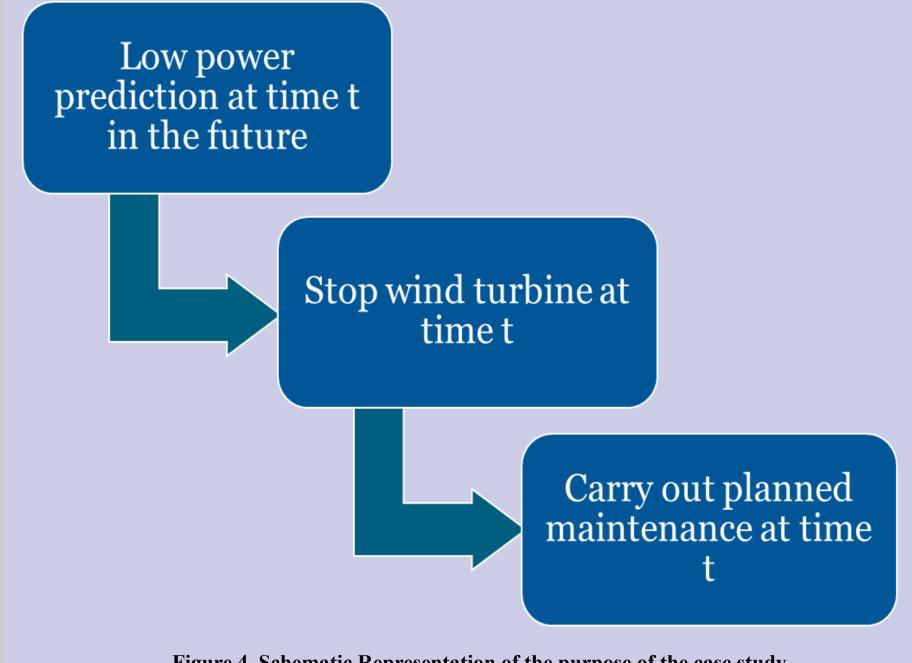


Figure 4. Schematic Representation of the purpose of the case study

# **Artificial Intelligence**

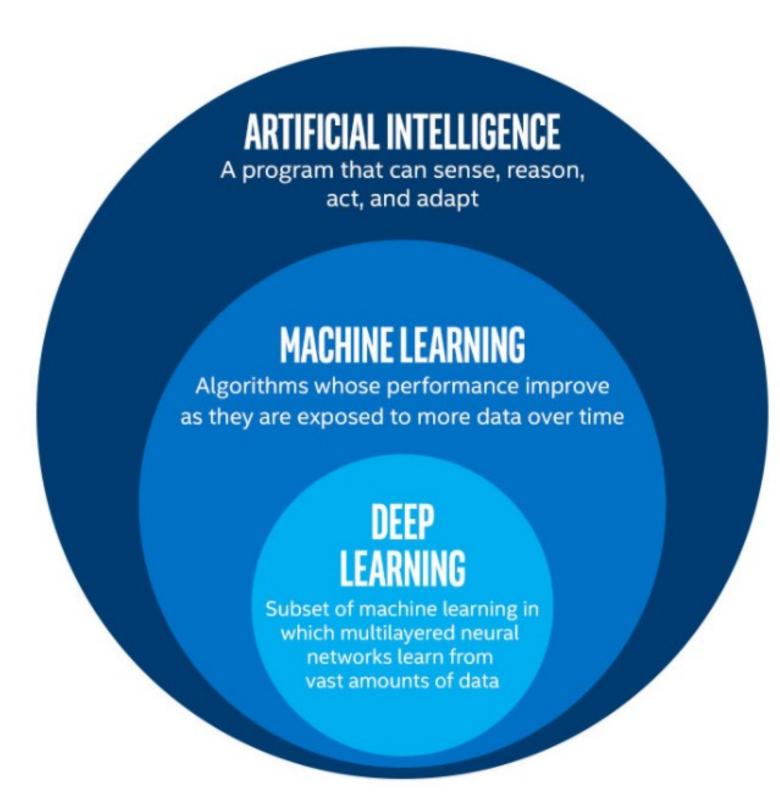


Figure 5. S.Srivastav (2020). Artificial Intelligence, Machine Learning, and Deep Learning. What's the Real Difference? Available at: https://medium.com/swlh/artificial-intelligence-machinelearning-and-deep-learning-whats-the-real-difference-94fe7e528097[Accessed on: 31Mar. 2022].

#### **Machine Learning**

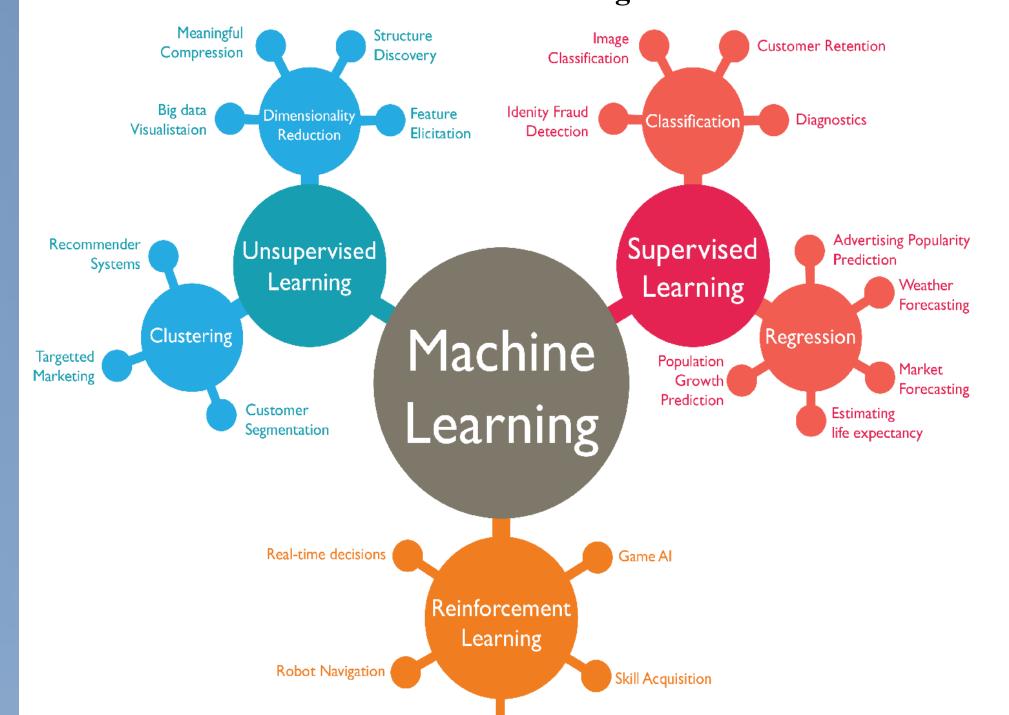
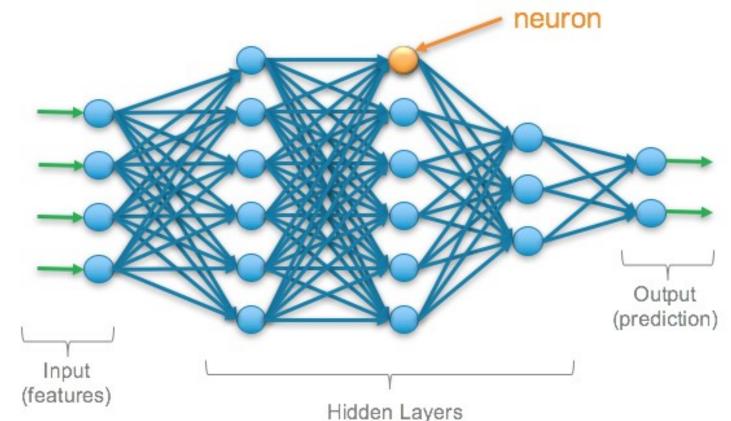


Figure 6. D.Shewan(2021).[Online]. Available:https://www.wordstream.com/blog/ ws/2017/07/28/machine-learning-applications. [Accessed: 30– Mar- 2022]

## **Deep Learning**



lots of layers ~ "deep learning" Figure 7. S.Ronaghan (2018).[Online]. Available at :https://srnghn.medium.com/introduction-to-deeplearning-what-do-i-need-to-know-75794ebc4a62[Accessed: 30– Mar- 2022]

#### Case Study 1 **Methodology-Preprocessing**

# Preprocessing of datasets and datapoints leftover at each stage

Dealing with uncompressed SCADA dataset

Initial number of datapoints: 28604042 After having identified that the power takes 1-2 mins to change trend accordingly after server status online (SecsOnline) changes from 0 to 1 the following step is done- After a change in servers status (SecsOnline), for the next 120 Secs ahead, all the datapoints are removed so as it does not influence the modelling part.

Datapoints left after this process: 23560125

•Remove negative power output -: 14102359

•Removing data points with brake conditions - 10745638 • Feature selection and importance using Random Forest Regressor -

selecting feature input for model •Resampling dataset to 4 mins to be able to use Mahalanobis distance for

Filtering SCADA outlier removal-131351 Drop NaN values -56192

Detection of outliers on power curve graph and removing them using Mahalanobis distance - 52970 Removing data points of power and wind speed which have a huge

**Outlier detection** and removals

dataset

amount of repetitive values which occur more than 20 times: 52970.(these are causing horizontal lines known as downtime on power curve)

Attributes

Comparison

- ·Visualisation of the offset between the trends of the same attributes recorded in the MET dataset and predicted in the METEO dataset ·Adjusting the values of the attributes in MET dataset to match same trends of attributes in the METEO dataset
- · Joining the dataset of SCADA with the MET dataset

•Datapoints left: 52970

•Dropping NaN values: 49208

Whole dataset treatment

•Feature scaling using normalization of inputs and outputs

#### Case Study 1 Model Development and Deployment **Model Deployment** Multiple METEO forecasts datasets combined into one single data frame at 1-hour interval ORE SCADA and MET ORE Average Wind **Deep Learning Model** Power output MET Average Wind predictions compared Artificial Neural Network Direction to actual power known as Multilayer MET Pressure Perceptron (MLP) output MET Temperature Label: SCADA Actual Power output Power predictions for up to 72 hours in the future at 1-

#### Case Study 1 **Results and Discussion**

hour interval time

Figure 9.Modelling method

The MLP model developed in this case study is compared to a baseline model developed and deployed on the same dataset. From the metric results, it can be seen that the combination of preprocessing techniques together with a deep learning algorithm work the best. The amount of data to train/test the model also improved the metric results.

I	Models	Inputs	Pre-processing techniques	Metric results
	The MLP Neural Network	Wind speed Wind direction Temperature Pressure	-Datapoints removed where there is a gap between server online change and changing power -Negative power removed -Brake conditions removed -METEO and MET trends attributes adjusted -Mahalanobis distance for outlier removal -Excessive repetitive values removed	R <sub>2</sub> : 0.98  MAE:194.14  RMSE:283.7  NRMSE:0.11  MAE between the power from SCADA and Power predicted using METEO forecasts:1462.16
r	Linear regression	Wind speed	-Datapoints removed where there is a gap between server online change and changing power -Negative power removed -Brake conditions removed -METEO and MET trends attributes adjusted -Mahalanobis distance for outlier removal -Excessive repetitive values removed	R <sub>2</sub> score:0.94 MAE: 441.1 RMSE:544.7 NMRSE:0.22  MAE between the power from SCADA and power predicted using METEO forecasts:2072.99
r	Linear regression	Wind Speed	None	R <sub>2</sub> :0.31 MAE:1202.9 RMSE:1663

Figure 10. Metric results of model developed compared to baseline model developed using same dataset

## **Maintenance Scheduling**

The energy output of the wind turbine in 4 hours blocks for up to 72 hours in the future is calculated so as to give enough time for the maintenance team to plan and do the maintenance needed.

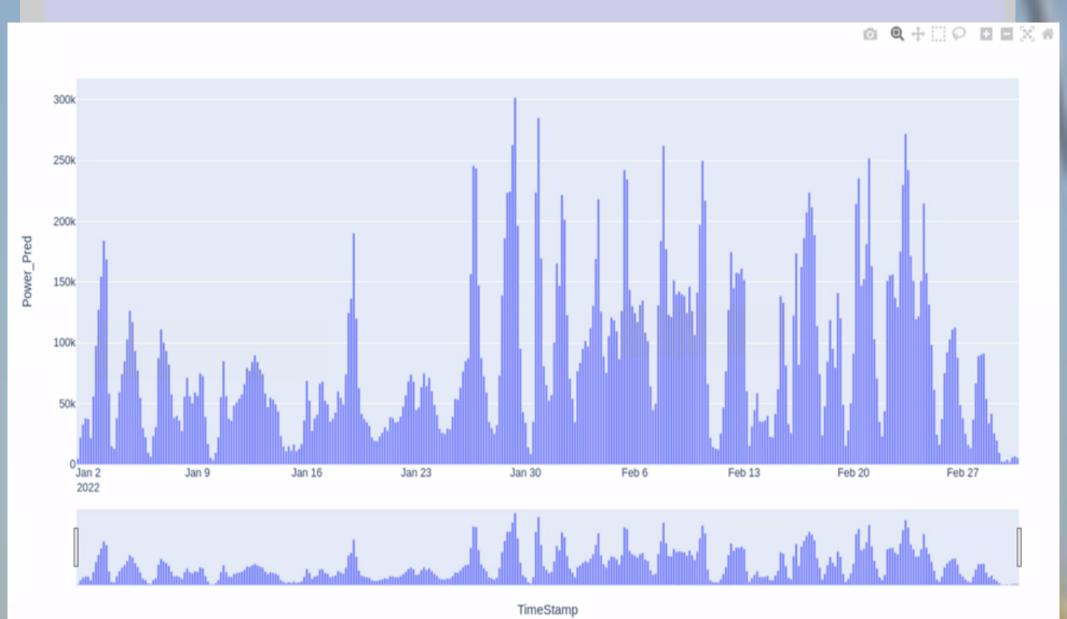


Figure 11. Energy output for up to 72 hours in the future in 4 hours bars

#### **Next Case Study** Optimisation of diesel usage in oil platforms

## **Current Status**

- Description and operation of oil platform
- Exploratory data analysis:
- ⇒ Features that affect usage of diesel
- ⇒ Label (output) attribute
- ⇒ Type of model to use



Figure 12. Ithaca Energy (2022)[Online]. Available at :https://www.ithacaenergy.com/our-operations/ ithaca-operated-assets/captain[Accessed: 30- Mar- 2022]

Figure 8 .Preprocessing techniques used in this case study