# Synoptic Project Outline

## Title

Using Machine learning to develop a proactive Auto-Scaler to improve energy efficiency of Kubernetes.

## Aim

Traditional Kubernetes auto-scalers automatically scale the size of pods, containers or nodes in a cluster based on the demands of the current workload. This means the demand for extra resources is already present before more can be provisioned and, by definition, makes them reactionary leading to inefficient resource utilisation and thus wastes energy.

This project aims to create a Machine Learning model that can be used by the auto-scaler to proactively scale resources in anticipation of demand. This will help avoid over provisioning whilst maintaining the required performance and availability, reducing the consumption of the system.

## Objectives

* Acquire a comprehensive understanding of Kubernetes and it’s auto-scalers, including its architecture, components, and deployment methods, by dedicating 10 hours a week to study through semester 1, aiming to achieve an understanding sufficient to deploy a basic Kubernetes cluster by week 11 of semester 1.
* Clean the acquired dataset, preparing it for use in training, testing, and validation of the machine learning model by the start of semester 2.
* Identify and acquire a suitable dataset for the Kubernetes resource utilisation within the next four weeks, resulting in a data source that accurately represent Kubernetes workloads.
* Develop a benchmark model to measure the current energy consumption of Kubernetes clusters within eight weeks to establish a baseline for evaluating the effectiveness of the proactive auto-scaler.
* Design and train a machine learning model to predict Kubernetes workload demand with to a degree of accuracy, utilising existing datasets, and complete this phase before week 6, Semester 2. That in turn reduces energy consumption.

## Deliverables

* A detailed report or presentation demonstrating a thorough understanding of Kubernetes and its auto-scalers, including a successfully deployed basic Kubernetes cluster.
* A fully cleaned and pre-processed dataset, ready for training, testing, and validation of the machine learning model.
* A secured dataset accurately representing Kubernetes workloads for resource utilization analysis.
* A developed benchmark model to measure the energy consumption of Kubernetes clusters, ready for baseline evaluations.
* A machine learning model, trained and tested to predict Kubernetes workload demand with at least 85% accuracy.
* A comprehensive Synoptic Project Report covering all objectives and aims of the project.

## Plan

### Semester 1

* Week 7-9
  + begin background information research.
  + Find database on a Kubernetes workload usage to train and validate the Machine Learning Model
  + Begin literature review.
* Week 9-11
  + Complete background research
  + Design Machine learning model and define the machine learning algorithm(s) to be used.
  + Establishing performance metrics and objectives.
  + Define model to quantify energy consumption from resource usage data.
* Before Semester 2
  + Have a completed background research and literature review chapter (1 & 2)
  + Have the Machine learning and energy consumption model defined.
  + Start Chapter 4 - Design

### Semester 2

* Week 1-2
  + Initial model building and testing on a small subset of data.
  + Complete Chapter 4, documenting the model's architecture, usage, and limitations.
* Week 2-4
  + Training the model on the full dataset. Monitoring for overfitting and making necessary adjustments.
  + Evaluating the model's performance against set metrics. Iterating on the model to improve accuracy and efficiency.
* Week 4-6
  + Start Chapter 5 - Implementation
  + Further model iteration
  + Testing the model in the auto-scaler to test compatibility and produce first set of real results.
  + Final tweaks and optimisations to the model
* Week 6-8
  + Complete the machine learning model.
  + Integrate the machine learning model into the auto-scaler and collect final usage data.
  + Cross reference new usage data with benchmark data to produce final performance and consumption results.
* Over Easter
  + Complete chapter 5 - Implementation
  + Start and Complete Chapter 6 - Evaluation
* Week 9-10
  + Add the finishing touches to the report.
  + Submit the report and all supporting documents and repositories.

## Risks

**Compliance with Benchmarks:**

* Ensuring the model's operations comply with industry standards in performance - I don’t want to prioritise energy efficiency so much that the performance benchmarks aren’t maintained. As it will

**Bias and Fairness:**

* Evaluating and mitigating any biases in the dataset to prevent unfair resource allocation by implementing concepts of Machine Learning and linear regression.
* Regularly assessing the model with validation datasets to ensure it’s

**Lack available Data**:

* Failing to find a suitable data set that provides enough data to satisfy training the machine learning model as well as finding a benchmark energy consumption figure; will require me to produce my own. This will add time onto my plan, but I have taken this into account and left myself the 4 weeks over Christmas free for this very issue - giving me time to set up, run and collect data of my own cluster’s usages.

## Ethics

**Data Privacy and Security**

* Ensuring data used to train the model does not violate user privacy (if it’s external data)

**Transparency**

* Ensure my models are openly available and all influence and utilisation of other methods and models are correctly cited.

**Accountability and Responsibility**

* Ensuring there is a mechanism for human intervention in case of errors or unintended consequences.