**EEX5362 Performance Modelling**

**Mini Project**

**Deliverable 01**

T. Yethum Danith

422514872

s22010334

**System Details (System Identification and Details)**

Complex System – **Cloud Workload Scheduling System**

The system that I selected is a cloud workload scheduling system.

This system is used in cloud computing platforms to manage how computer resources are shared among users and tasks.

In this system, many users do tasks with cloud. For example, we can say uploading files, running programs and processing data. So, the system must decide which system should be executed first. Not only that, but it should also decide how much CPU and memory each task required. And system should handle multiple users at once.

If the scheduling is poor, tasks may take longer and because of that servers may overload, or system may slow down.

**Dataset**:

The dataset contains 5000 records of different tasks that run in a cloud system. Each record includes details about how the system performed for one task.

The database contains columns like task start time, task end time, task waiting time, CPU utilization, Number of active users and so on.

So, this dataset is useful, because from this data we can measure and see how the system performance changes under different workloads, users and priority.

**Performance Objectives**

The main purpose of this is to understand how the cloud workload systems perform under different conditions. And from that we can find ways to make it more efficient. By focusing on performance metrices, we can identify where time and resources are being wasted and what improvements can be made to the system to make it faster and more reliable.

* Reduce task waiting time

When lots of users send tasks to the cloud system, some tasks have to wait in a queue before they start to run. A long waiting time means the system is slow to respond.

* Increasing System Throughout

This measures how many tasks can be completed per second in the system in different conditions and workloads. And it should increase the throughout without decreasing the quality.

* Use resources efficiently

The CPU, memory and network are limited resources. So, there should be balanced utilization for all. All the resources should be efficient without being overloaded.

* Reduce Taks Execution Time

By improving scheduling and reducing delays from the resource sharing, we can reduce the task execution time. Then we can make the system faster and improve user satisfaction.

* Error Rate

Here we focus on error rates because there can be lots of tasks that due to that system may fail or stop. Then it will be considered as an error. So, the goal is to reduce the error rate.