Koç University
College of Engineering
INDR 491 Engineering Design Project Proposal

Fall 2023

Design of an Analytical Framework for Data Center Efficient Energy Management

Information about the team:

Name	ID	Email	Phone
Umur Berkay Karakaş	0069075	ukarakas18@ku.edu.tr	+905423369770
Serra Işık	0076808	sisik20@ku.edu.tr	+905330398208
Bora Aktaş	0071461	baktas19@ku.edu.tr	+905426845683
Batuhan Gül	0071780	bgul19@ku.edu.tr	+905439166782
Sena Karadağ	0071880	skaradag19@ku.edu.tr	+905531603161

Academic Supervisor:

Name	Signature
Fikri Karaesmen	

Abstract

KoçSistem's data center consists of multiple rooms, each with different layouts, a cooling system, and various rack units. The current cooling system uses water-based chiller systems and Direct Expansion (DX), and their parameters are determined based on experience rather than automation. This project aims to develop models to help reduce KoçSistem's high energy consumption in its server rooms. We intend to prevent overheating on servers by optimizing the use of cooling systems. Key aspects of our project include sensitivity analysis to understand the effect of different variables on energy use and PUE value. The primary approach will be to use machine learning models to analyze historical data, predict PUE, and optimize. By analyzing the past data, the objective is to reduce energy usage, and provide automated decision-making. For this project, tools like Jupyter Notebook in Python will be employed, along with libraries such as NumPy, SciPy, torch, Pandas, LazyPredict and scikit-learn. In conclusion, the project aims to offer a more efficient, cost-effective solution to KoçSistem's data center cooling challenges.

Table of contents

Section 1 System Description

KoçSistem is a well-known company that was founded in 1945 and has its main office in Istanbul. It is part of a bigger group called Koç Bilgi Grubu AS. This company is known for providing services in the area of information technology. Over the years, KoçSistem has expanded its services and also provides network solutions to its customers. KoçSistem was earlier known by the name Koç-Unisys before it was renamed.

The company has expanded and offers more services in the field of information technology. Now, it provides consulting services and helps other companies integrate IT systems. It also creates software products for their customers. KoçSistem is not only known in Turkey but also in more than thirty other countries, especially in Azerbaijan, Malaysia, England, and China.

KoçSistem offers a wide range of services. Some of these services include solutions related to the Internet of Things (IoT), analytics to help businesses understand data better, business solutions to and data center and cloud services for storing and managing data.

KoçSistem has its own data center in Üsküdar, Istanbul. It has 2 floors and consists of 12 different rooms that store server racks. Some of the rooms are customized to store servers dedicated to single servers, whereas other rooms consist of server racks of multiple servers. The layout of the rooms are not identical, and the building where their data center is located was built as a hospital. The floors where the data center is located was meant to be the mortuary.

Section 2 Problem Definition

2.1 *Analysis of the current system*

The existing setup includes many rooms with different layouts and cooling systems. Each room contains lots of rack units, and the purpose of the data center is to create a sufficiently cold environment to prevent extra heating and humidity, which is harmful for servers.

To keep each room in the data center at a consistent temperature, the existing cooling system makes use of room-based cooling systems. The two primary components of the system are water-based chiller systems and Direct Expansion (DX). There are different cooling needs for varying rack units with different brands and designs. In addition, physical layout challenges, such as inaccessible columns, also prevent cool air from being placed efficiently. Changes in the outside temperature have a big impact on how much cooling is needed inside as well.

The majority of cooling system decisions are made according to the experience, therefore the decision-making process is not automated. There are variables called "set parameters" for temperature and humidity that indicate the optimal values for those variables in the room. It is expected that these set parameters will remain within the defined boundaries. The set parameter values are determined based on the current season. Furthermore, the operation cycles (generally turning off and on) of the cooling units are generally determined on a weekly basis through manual decision-making.

The current system's Power Usage Efficiency (PUE) values fluctuate between 1.3 and 2.0. These values are considered low when compared with the other data centers in Turkey. It is important to remember that these improvements in recent years have reduced PUE values and led to better energy preservation.

There is room for improvement in the current data center cooling system through the use of automated decision-making and historical data analysis by machine learning models.

2.2 *Symptoms and complaints*

Our recent evaluation of the data center has highlighted several issues that require attention:

- **Overcooling**: The temperature of the data center is lowered more than necessary, causing significant energy wastage.

- **Different types of Rack Units:** There are lots of rack units with different types that are the reason for the different cooling requirements needed. However, since we have a room-based cooling system, we should run a system that meets the needs of all of the types.

- **Diverse Cooling Systems:** The use of different types of cooling systems, such as air-based (DX) and water-based (Chillers), complicates the uniform management of cooling.

- **Cooling Units Placement:** The current positioning of the cooling units might not be optimal for cooling.

- **Rack Unit Placement:** The organization and positioning of the rack units may be leading to cooling issues, possibly resulting in turbulence in the aisles. There are also certain locations that are difficult to access, especially behind the building's columns, which makes it inefficient because it hinders the cold air from reaching there.

- **Humidity Concerns:** Excessive humidity can be harmful to the equipment, leading to hardware malfunctions or decreased lifespan.

- **External Temperature Influence:** Unpredictable changes in the outside temperature seem to have an evident effect on the internal cooling system.

- **IT Load Fluctuations:** Changes in the IT load on the servers result in temperature spikes, requiring the cooling system to operate more intensively. If IT load is not evenly distributed, it causes distinct heating zones within the room

- **Server Overheating:** If the server's temperature exceeds the threshold, it gives error and shuts itself down for self-protection, which is a highly undesirable situation. # yeri

- **Delays for Desired Temperature:** Even when the cooling systems start running, it takes some time for the room temperature and servers to cool down to the desired level.

2.3 *Problem definition and scope*

KoçSistem is currently striving with an inefficient cooling scheduling system across its data centers. This inefficiency is caused by running cooling systems more than required. Therefore, the Power Usage Effectiveness (PUE) value, which depends on the change in total server IT load, temperature, humidity and energy consumption of cooling systems, is higher than desired.

One of the biggest obstacles is the lack of predictability regarding server workload. This leads to manual and often imprecise cooling operations. This uncertainty causes strain on the server and disruption of operations

The high PUE value proves inefficiency in energy consumption. The problem is that cooling operations arrangements are performed manually because there is a lack of predictability regarding server workload. This uncertainty causes strain on the server and disruption of operations. Consequently, this escalates the IT load on the data servers, reduces job performance and potentially triggers customer dissatisfaction that could financially penalize KoçSistem through fines.

The primary objective of this project is to optimize the scheduling of the cooling system in KoçSistem's data centers. The project does not include any modifications to the existing physical air conditioning systems. Instead, it emphasizes the intelligent scheduling of cooling operations to reduce energy consumption and minimize cooling expenses.

2.4 *Objectives of the project*

Considering the symptoms and the problem explained in the problem definitions section, the main aim of this project is to create a mathematical model in order to help solve KoçSistem's high energy consumption and cut cooling costs of their server rooms in the data centers.

Once the project is complete, we will improve server efficiency by stopping the CPUs from overheating and consequently reducing any customer job delays. We aim to develop a mathematical model to optimize the usage of cooling systems in the server rooms. With the developed model, we are planning to make employees effectively manage energy usage. The objective of this approach is to efficiently determine the most suitable cooling scheduling and times in response to the workload of servers, including both planned and unplanned server loads. The use of this dynamic technique provides optimal functioning of the data servers while minimizing energy usage.

A significant part of our project is sensitivity analysis. With sensitivity analysis we will check how different things, like the amount of workload, or changes in outside temperature or how much a cooling system is being utilized, can affect the energy use and PUE value. Also, this will allow us to keep the server temperature within a certain range by observing the effects of such variables on server temperature.

2.5 *Literature review*

There have been various approaches to make data centers more efficient using both predictive and optimization tools. Prediction-based approaches mainly focus on predicting future efficiency metric values, most importantly Power Usage Efficiency (PUE), based on historical data provided by the physical sensors in the server room. Optimization-based approaches include linear and dynamic programming on both hardware and software aspects of data center frameworks.

2.5.1 *Prediction-based Approaches*

The first ground-breaking predictive method for data center efficiency was proposed by Gao (2014), from Google. His main goal was to create a model that predicts PUE from Google data centers' 2-year historical data using MLP (multilayer perceptron) with 5 hidden layers and 50 nodes in each hidden layer. The features of the neural network were as follows:

1. Total server IT load [kW]
2. Total Campus Core Network Room (CCNR) IT load [kW]
3. Total number of process water pumps (PWP) running
4. Mean PWP variable frequency drive (VFD) speed [%]
5. Total number of condenser water pumps (CWP) running
6. Mean CWP variable frequency drive (VFD) speed [%]
7. Total number of cooling towers running
8. Mean cooling tower leaving water temperature (LWT) setpoint [F]

9. Total number of chillers running
10. Total number of drycoolers running
11. Total number of chilled water injection pumps running
12. Mean chilled water injection pump setpoint temperature [F]
13. Mean heat exchanger approach temperature [F]
14. Outside air wet bulb (WB) temperature [F]
15. Outside air dry bulb (DB) temperature [F]
16. Outside air enthalpy [kJ/kg]
17. Outside air relative humidity (RH) [%]
18. Outdoor wind speed [mph]
19. Outdoor wind direction [deg]

Having such a robust and reliable model allowed Google to do sensitivity analysis and anomaly detection easily. Using controllable variables from the listed features, they analyzed their effects on PUE on several scenarios to decrease PUE.

$$PUE = \frac{P_{total}}{P_{IT}} = \frac{P_{cooling} + P_{IT} + P_{electricalLosses} + P_{misc}}{P_{IT}}$$

For an efficient data center, the power consumption for non-IT components should be as small as possible, i.e., PUE should be close to 1. In anomaly detection scenarios, they used the MLP model to catch erroneous readings of the individual sensors when actual PUE values were higher than predicted PUE values.

Shoukourian et al. (2017) proposed to model the Coefficient of Performance (COP) rather than PUE, as they discussed that COP indicates the power consumption of four cooling circuits, whereas PUE is the combination of COPs from all cooling elements in one number.

$$COP = \frac{Q_{CoolingCircuits}}{P_{CoolingCircuits}}$$

where $Q_{CoolingCircuits}$ is the aggregated amount of cold generated by four cooling circuits (in watts) and $P_{CoolingCircuits}$ is the aggregated amount of power generated by four cooling circuits (in watts).

They argued that using COPs, it is easier to observe the effects of various control variables on the overall efficiency. In their model, they used LSTM with two hidden layers and a final linear layer with the following features extracted from Leibniz Supercomputing Centre (LRZ):

1. Aggregated amount of cold generated by each cooling circuit
2. Aggregated amount of power consumed by the fans of each cooling tower
3. The number of active cooling towers
4. Wet bulb temperature
5. Inlet water temperature (to the distribution bar) from each cooling circuit
6. Return water temperature (to the distribution bar) to each cooling circuit

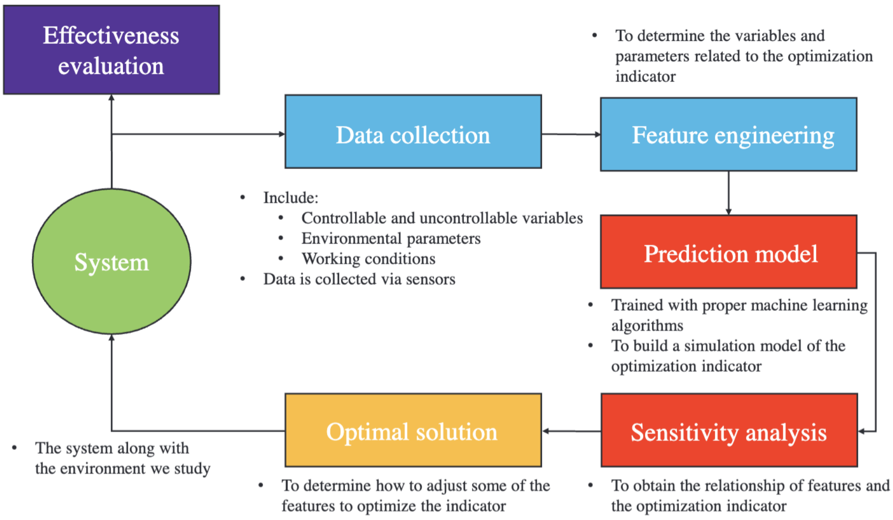


Figure 1: PUE optimization framework of Yang et. al (2021)

Yang et al. (2021) proposed another approach similar to Gao (2014) for PUE optimization framework. First, they collected data from individual sensors in Tencent Tianjin Data Center building No. 4, containing 38 features, similar to the features Gao (2014) used, ranging from January 2018 until November 2019. After applying some feature engineering, they set up different prediction models, i.e., neural networks (NNs), light gradient boosting machine (LightGBM), random forest and RNN. Finally, they conducted sensitivity analysis within the constraints to optimize PUE.

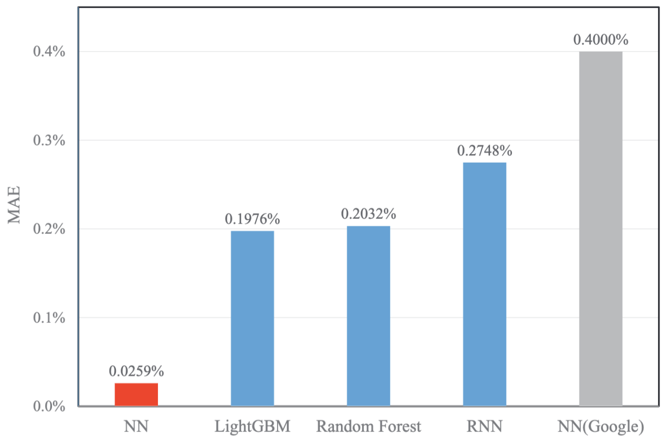


Figure 2: MAE of models Yang et al. (2021) and Gao (2014) used

They found out that 10 most important features affecting PUE are total server IT load, condenser water flow of chillers, outside air enthalpy, chillers current percentage, condenser water input temperature of chillers, chilled water output temperature of chillers, indoor air enthalpy, condenser water output temperature of cooling towers, chilled water of chillers and outside air wet bulb temperature. Among the controllable variables, the condenser water flow of chillers is the most important variable. They also outperformed Gao (2014) in PUE prediction model and their best performing model was a neural network with 3 hidden layers.

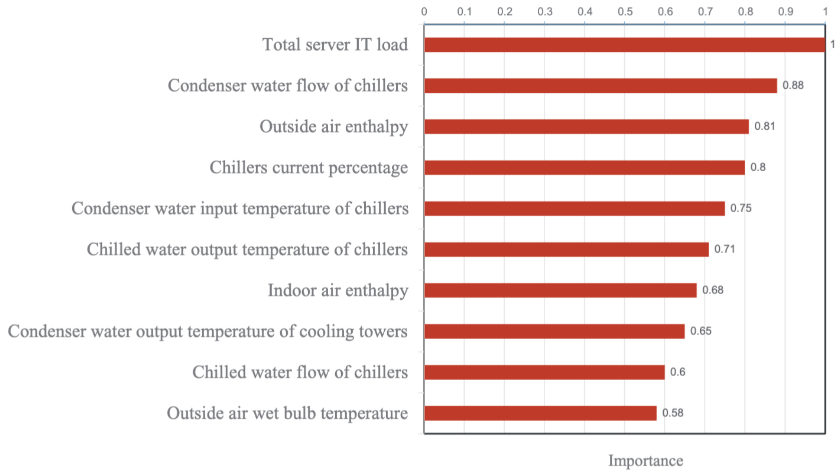


Figure 3: Top 10 most important features affecting PUE from Yang et al. (2021)

2.5.2 *Optimization-based Approaches*

Two optimization based approaches that focus on different aspects of data center systems have previously been proposed. Stamatescu et. al (2018) came up with an approach in which they optimize the scheduling of virtual machines (VMs) in the data center servers with a mixed integer linear programming problem. The decision variables are the assignment of each virtual machine to each server at time t. The constraints mainly focus on the CPU and memory capacity of each server. The objective of the problem is to minimize the energy consumption caused by working servers, migration of VMs between different servers and turning servers on/off in consecutive time steps.

COMMERCIAL ALGORITHM - ENERGY CONSUMPTION

Hour	1	2	3	4	5
Server 1	660	632	445	335	529
Server 2	160	166	158	146	135
Server 3	142	180	180	178	173
Server 4	10	174	269	8	0
Server 5	5	244	250	272	7
Server Consumption	977	1396	1302	939	844
Total IT Consumption of the Data Center: 5458 kWh/hour					

THE PROPOSED ALGORITHM - ENERGY CONSUMPTION

Hour	1	2	3	4	5
Server 1	628	740	645	0	0
Server 2	0	0	0	140	0
Server 3	0	0	0	0	0
Server 4	0	250	261	250	25
Server 5	258	270	261	270	265
Server Consumption	886	1260	1167	660	290
Total IT Consumption of the Data Center: 4263 kWh/hour					

Figure 4 and 5: Energy consumption comparison between commercial algorithm and Stamatescu et. al (2018)'s algorithm

In a comparison with a commercial algorithm, their approach decreased the total IT consumption by 22%, in a system with 5 servers and 5 types of application.

Lazic et al. (2018) comes up with a linear optimization problem with a quadratic objective, by which they optimize the scheduling of air handling units (AHUs) to control the temperature and the airflow of the data center.

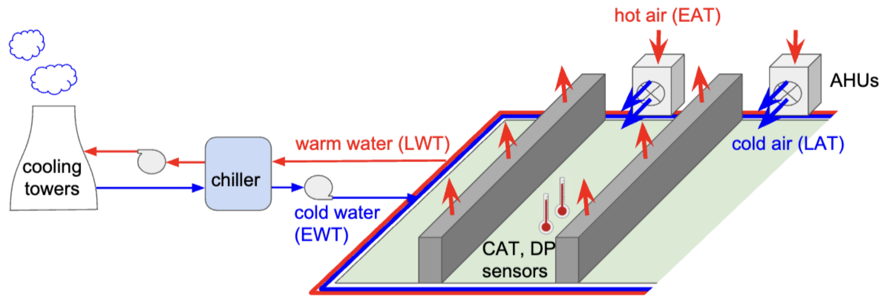


Figure 6: Data center cooling loop from Lazic et al. (2018)

They classified the variables of a data center system into three sections. “Control” variables (variables that can be manipulated) of their system were “fan speed” and “valve opening” for each AHU. “State” (variables that are predicted and regulated) variables were “differential air pressure” (DP), “cold-aisle temperature” (CAT), “entering air temperature” (EAT), and “leaving air temperature” (LAT). “Disturbance” variables were “server power usage” and “entering water temperature” of the chilled water at each AHU. Their optimization model had 1.2K variables along with a “large number” of linear and range constraints.

2.6 *Methodology*

Our methodology will be mainly defined by the data that we will either be provided by KoçSistem or find on the Internet. Considering the previous prediction-based research for data center efficiency, we are planning to implement a PUE optimization framework like Gao (2014) and Yang et al. (2021) performed. We will perform sensitivity and feature importance analysis in order to understand the effects of control variables of a DC system on PUE. Combining the effects of multiple control variables, we are planning to find an optimal solution at each timestep under different state variables.

For the PUE optimization framework, we will first train a machine learning model with the historical data that predicts PUE in each timestep. In the optimization step, PUE would be minimized by shutting down all cooling systems. Therefore, we should take constraints into consideration as well. During our visit to KoçSistem, we learned that “leaving air temperature” in AHUs should 24 ℃ ∓3. For such constraints, we will train additional models to learn the relations between state & control variables and such variables. Our goal in training such models is that during optimization, we will predict those variables in the constraints using optimized control variables so that we can satisfy those constraints to meet system requirements.

In terms of a mathematical model, our primitive model would be as follows:

$$
\begin{array}{lll}
\text{minimize} & \alpha * PUE & \\
\text{s.t.} & P(V) = PUE & \\
 & C_i(V) = c_i & \text{for } \forall i \in C \\
 & c_i \le c_{i_{max}} & \text{for } \forall i \in C \\
 & c_i \ge c_{i_{min}} & \text{for } \forall i \in C
\end{array}
$$

where P is the model that predicts P, V is the set of all variables in the data, C_i is the model that predicts constraining variable c_i using all variables and c_{i_max} and c_{i_min} are the corresponding bounds of constraining variable c_i.

Section 3 Project Outcome and Deliverables

3.1 *Expected outcomes (deliverables) of the project*

This project's primary goal is to establish a decision-making model that manages the cooling schedule of the KoçSistem Data Center server rooms. With the support of critical data such as anticipated IT load and upcoming weather conditions within the planning horizon, this system is programmed to be activated once the previous schedule concludes, thereby deriving the subsequent cooling schedule for the following time period. One of the most important KPIs is the Power Usage Effectiveness (PUE), which is a metric that represents the ratio of the total power consumption of the data center to the IT-related power consumption. PUE is currently between 1.3 and 2 in KoçSistem data center, which is the same as the average in Europe. However, through this project, our goal is to decrease the PUE significantly while maintaining the server temperatures and humidity at desired levels.

In addition to decreasing energy consumption and PUE, our model will enable KoçSistem to have automated decision making for cooling scheduling. Having such a model, there would be less need to physically observe the server conditions in case of an emergency. Also, in case of a loss of an experienced worker responsible for cooling in the data center, the cooling process would not be hindered thanks to our model.

We hope that our model would be implemented to KoçSistem's current data server cooling systems. However, because of the many limitations of the current system and the assumptions that we made, we believe that our model will not be directly implemented. Even though that is the case, we will report the estimated improvement on PUE and the effect of each decision variable on the system according to the data provided to us.

Some other benefits also result from the implementation of our model for scheduling cooling systems. In addition to reducing energy consumption and resulting in cost savings, application of such a model also reduces carbon emissions, therefore contributing to a more environmentally responsible operational model. Considering the total energy consumption of data centers worldwide, optimization of energy consumptions in data centers could be important for the environment and might be a positive step against the current climate change problem.

3.2 *Major limitations, assumptions and risk analysis*

3.2.1 *Limitations*

The current layout design of the data center is inflexible to changes. This prevents the optimal arrangement of rack units and the positioning of cooling systems. There are diverse rack units, each with individual needs, which force us to take a holistic approach when considering the cooling of rack units. One limitation is the current hardware quality. Upgrading the machines, which could significantly enhance efficiency, is not an option in the existing system.

3.2.2 *Assumptions*

We will assume that when a cooling system is activated, there's a specific duration required to stabilize the room at the desired temperature.

3.2.3 *Risks*

The most significant risk is that servers might shut themselves down if they detect excessive heat, leading to potential data loss or service interruptions. To prevent this, our strategy will be to maintain temperatures within a secure range, ensuring sudden temperature spikes don't reach critical levels.

3.3 *Impact*

As a result of this project, by reducing unnecessary cooling, energy consumption will be reduced, resulting in less electricity costs. The decision-making process will be transformed from manual to a more automated system, preventing unreliable system decisions. Our model will predict the effectiveness of any given decision on the system and how it can potentially lower costs.

Furthermore, by analyzing the influence of cooling systems on temperature, we can use the machines more effectively, ensuring optimal usage and more energy savings. Our project aims to reduce the inefficiencies in the current system, leading to both operational and financial improvements for the company.

3.4 *Resources*

For our research, we have used Google Scholar for finding academic papers in the relevant area. We will be mainly utilizing Jupyter Notebook in Python language. For machine learning purposes, we will use relevant Python libraries such as NumPy, SciPy, torch and scikit-learn. In case there will be a need for web scraping for data collection, we will use Python libraries such as BeautifulSoup, requests and Selenium. For internal communication and file sharing, we will use Notion and Google Drive. In the case that we are not provided any data by KoçSistem regarding their data center, we will use the dataset from Le (2021).

Section 4 Project Plan

4.1 *Work Package Descriptions*

Work package number	**1**	Start date or starting event:		**16.10.2023**	
Work package title	**Background Research and Analysis of the System**				
Participant number	**1**	**2**	**3**	**4**	**5**
Participant name	Batuhan Gül	Bora Aktaş	Sena Karadağ	Serra Işık	Umur Berkay Karakaş
Weeks per participant	2	2	2	2	2

Objectives

- Understanding operations and constraints in KoçSistem Data Center.
- Detecting similar topics that are discussed in KoçSistem Data Center and literature.
- Identifying gaps to optimize in KoçSistem Data Center.

Description of work

T1.1 (16.10.2023 – 23.10.2023) Literature Review

Doing literature review on data center cooling systems before visiting the data center.

T1.2 (23.10.2023 – 24.10.2023) Observing KoçSistem Data Center

Visiting the KoçSistem data center and understanding the underlying operations. Understanding the variables, constraints and details of the system.

T1.3 (24.10.2023 – 27.10.2023) Deciding Project Objectives, Scopes, and Methodology

Brainstorming among team members to decide the objective and the methodology according to the constraints and limitations of the system.

Deliverables

D1.1 (27.10.2023) Project Proposal

D1.2 (27.10.2023) The first sights that we made about the system can be delivered.

Milestones

M1.1 (24.10.2023) Operations and constraints are detected.

Work package number	**2**	Start date or starting event:	**27.10.2023**	
Work package title	**Data Collection and Preprocessing**			
Participant number	**1**	**2**	**3**	**4**	**5**
Participant name	Batuhan Gül	Bora Aktaş	Sena Karadağ	Serra Işık	Umur Berkay Karakaş
Weeks per participant	1	1	1	1	1

Objectives

- Getting the related data from KoçSistem Data Center.
- Detecting outliers in the data.
- Understanding the features in the data and making first observations

Description of work

T2.1 (27.10.2023 – 30.10.2023) Data Collection

Obtaining the data from KoçSistem or generating synthetic data or finding open-source data center cooling data.

T2.2 (30.10.2023 – 02.11.2023) Data Preprocessing

Inspecting the data and detecting outliers, deciding which outliers to be kept or removed.

T2.3 (02.11.2023 – 03.11.2023) Understanding the Current Features

Investigating the connections between features in the data, finding some patterns and correlations.

Deliverables

D2.1 (30.10.2023) Outlier-free data
D2.2 (03.11.2023) Patterns in the data that we detect in the first inspections

Milestones

M2.1 (30.10.2023) Real or synthetic data is obtained.
M2.2 (30.10.2023) The data is clearly analyzed, and it is matched with the real world processes in the data center.

Work package number	3	Start date or starting event:		03.11.2023	
Work package title	**Data Analysis and Feature Engineering**				
Participant number	**1**	**2**	**3**	**4**	**5**
Participant name	Batuhan Gül	Bora Aktaş	Sena Karadağ	Serra Işık	Umur Berkay Karakaş
Weeks per participant	1	2	1	2	1

Objectives

- Adding some features that we think can affect prediction power of the models that will be developed
- Analyzing the feature importance and deciding which features should stay in the models.

Description of work

T3.1 (03.11.2023 – 08.11.2023) Feature Engineering

Adding extra features such as days, seasons, and predicted weather information that meteorology published, or the important days in Turkey (like national holidays, Christmas, and so on)

T3.2 (08.11.2023 – 10.11.2023) Data Analysis

Investigation of the importance of the features and correlations.

T3.3 (10.11.2023 – 17.11.2023) EDA and Visualization

Visualization of the data and conducting exploratory data analysis. Presentation of our findings more intuitively.

Deliverables

D3.1 (10.11.2023) Feature importance
D3.2 (17.11.2023) Visualized data plots

Milestones

M3.1 (10.11.2023) The decision of which features will remain is made.
M3.2 (17.11.2023) Correlation of data is reported as visualized plots.

Work package number	4	Start date or starting event:		**10.11.2023**	
Work package title	**Developing Predictive Models**				
Participant number	**1**	**2**	**3**		
Participant name	Bora Aktaş	Serra Işık	Umur Berkay Karakaş		
Weeks per participant	3	2	3		

Objectives

- Design a predictive model structure for the data that has been analyzed.
- Develop the predictive model by utilizing best practices.
- Evaluate the model and interpret the results.

Description of work

T4.1 (11.11.2023 – 16.11.2023) Design of The Models

Exploration of some basic predictive models (linear regression, naive methods) and advanced machine learning models (KNN, neural networks, random forests). Using a Python library called "LazyPredict" to explore more predictive models that we do not know but can be suitable for the data. Defining our parameters and constraints and deciding where to split for training.

T4.2 (16.11.2023 – 27.11.2023) Developing the Predictive Model

Initiating the actual development phase. Training the model, optimizing it, and using regularization if overfitting occurs.

T4.3 (16.11.2023 – 27.11.2023) Training and Testing

Evaluation of the models performance in training and testing data separately using metrics such as RMSE, MAE, etc. Interpretation of the results of each model and understanding the advantages and disadvantages.

T4.4 (27.11.2023 – 01.12.2023) Model Selection

Elimination of the models that are evaluated in the training and testing phases according to their performances.

Deliverables

D4.1 (27.10.2023) Complete predictive models and their performances in the data

Milestones

M4.1 (24.10.2023) Reliable predictive models approved by supervisor
M4.2 (24.10.2023) In light of the data size, good MAE and RMSE values will be found.

Work package number	5	Start date or starting event:	01.12.2023	
Work package title	**Sensitivity and Risk Analysis**			

Participant number	**1**	**2**	**3**	**4**	**5**
Participant name	Batuhan Gül	Bora Aktaş	Sena Karadağ	Serra Işık	Umur Berkay Karakaş
Weeks per participant	1	1	1	1	1

Objectives

- Observing the effect of each decision variable on PUE of the system and server temperatures, both individually and jointly
- Determining the bounds of control variables which would satisfy system constraints such as set parameters of temperature and humidity.

Description of work

T5.1 (01.12.2023 – 05.12.2023) Sensitivity Analysis of Decision Variables

Determination of the parameters that have a greater effect on PUE and server temperature, and prioritization in the sensitivity analysis accordingly.

T5.2 (05.12.2023 – 08.12.2023) Risk Analysis for the Set Parameters

Examination of the boundaries of control variables which affect the set parameters of the system.

Deliverables

D5.1 (08.12.2023) The correlation between the PUE of the system and each decision that we are making while managing the cooling system of the data center

Milestones

M5.1 (08.12.2023) Sensitivity analysis plots for various control variables

Work package number	**6**	Start date or starting event:	**10.11.2023**		
Work package title	**Mathematical Modeling for Decision-Making**				
Participant number	**1**	**2**	**3**		
Participant name	Bora Aktaş	Serra Işık	Umur Berkay Karakaş		
Weeks per participant	3	4	4		

Objectives

- Formulate a mathematical model that represents the real world
- While considering constraints and objectives, implementing the mathematical model
- Analyzing the results of the model

Description of work

T6.1 (10.11.2023 – 24.11.2023) <u>Designing the Mathematical Model</u>

Identifying the problem, constraints, and objectives and using mixed-integer linear or nonlinear models to create a representation of the problem

T6.2 (24.11.2023 – 01.12.2023) <u>Developing the Model</u>

The model is implemented in a Python environment, using Python libraries. Testing the developed model simultaneously and adjusting the model parameters accordingly.

T6.3 (01.12.2023 – 08.12.2023) <u>Results and Decision Making</u>

The model will be tested on the test data, and it will be compared with the results of the decisions already taken by the KoçSistem without considering the future. These outcomes will be analyzed, and key findings will be determined.

Deliverables

D6.1 (27.10.2023) A Mathematical Model
D6.2 (27.10.2023) Informed decisions and recommended changes for the current system

Milestones

M6.1 (24.10.2023) A reliable mathematical model, that considers all of the constraints, is developed for further use.
M6.2 (24.10.2023) Automated decision-making model

<table>
<tr><td>Work package number</td><td>7</td><td colspan="2">Start date or starting event:</td><td colspan="2">16.10.2023</td></tr>
<tr><td>Work package title</td><td colspan="5">Reporting</td></tr>
<tr><td>Participant number</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td></tr>
<tr><td>Participant name</td><td>Batuhan Gül</td><td>Bora Aktaş</td><td>Sena Karadağ</td><td>Serra Işık</td><td>Umur Berkay Karakaş</td></tr>
<tr><td>Weeks per participant</td><td>14</td><td>5</td><td>14</td><td>5</td><td>5</td></tr>
</table>

Objectives

- In formal report, present our findings, insights and recommendations
- Provide comprehensive explanation of the project, including methodologies and results
- Communicate with the company

Description of work

T7.1 (30.10.2023 – 19.01.2024) Drafting the reports

Documentation of our findings, visualizations, and suggestions through the semester

T7.2 (13.01.2024 – 19.01.2024) Review and Final Report

Reviewing the draft for inconsistencies and finalizing the report according to the feedback from our advisor, company, and team members.

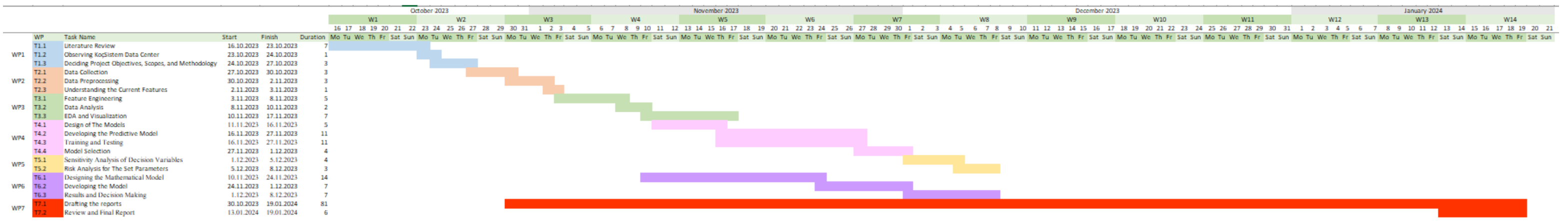
Deliverables

D7.1 (08.12.2023) Progress Report
D7.2 (12.01.2024) Final Poster
D7.3 (19.01.2024) Project Report

Milestones

M7.1 (19.01.2023) Finalization of the project
M7.2 (19.01.2023) Presentation of the project

4.2 *Gantt Chart*

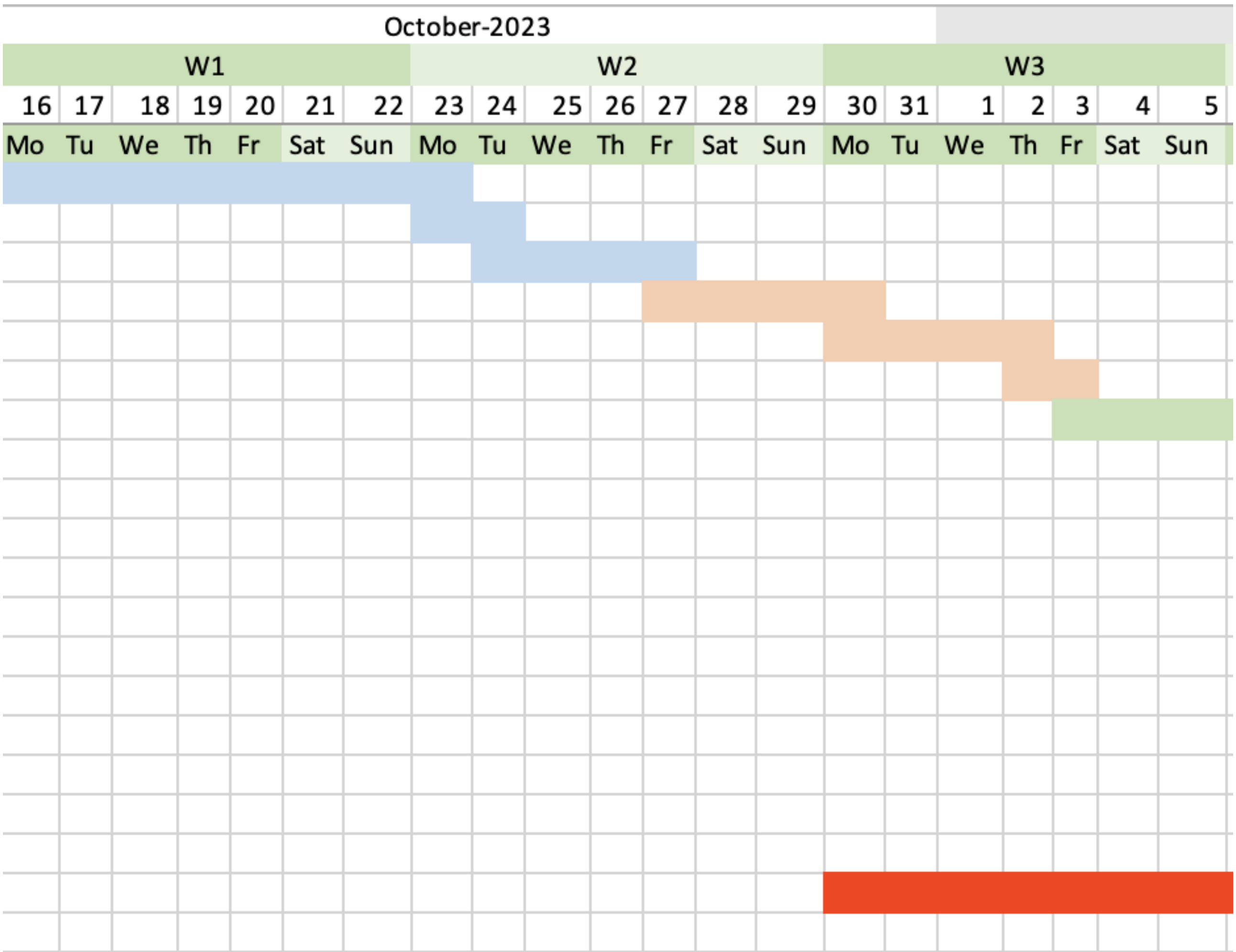


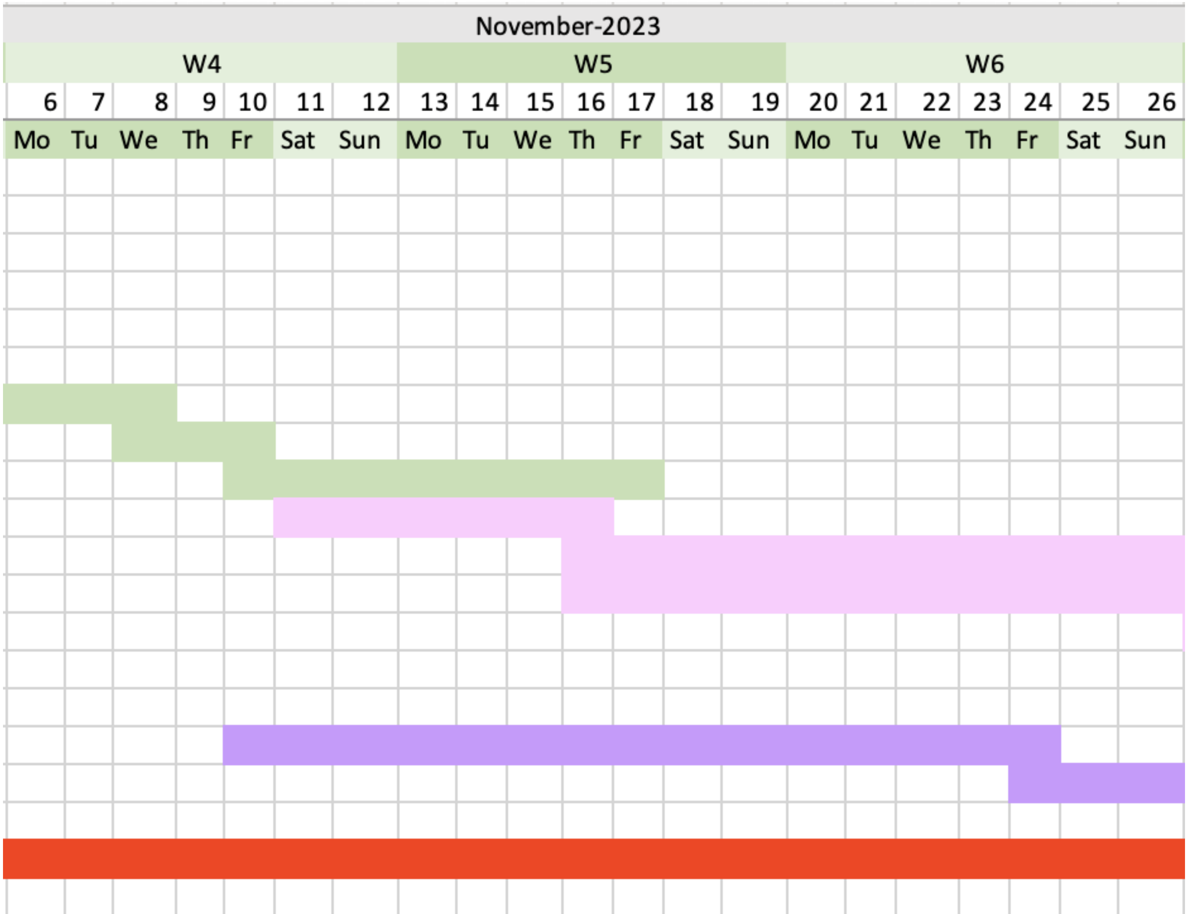
Section 5 References

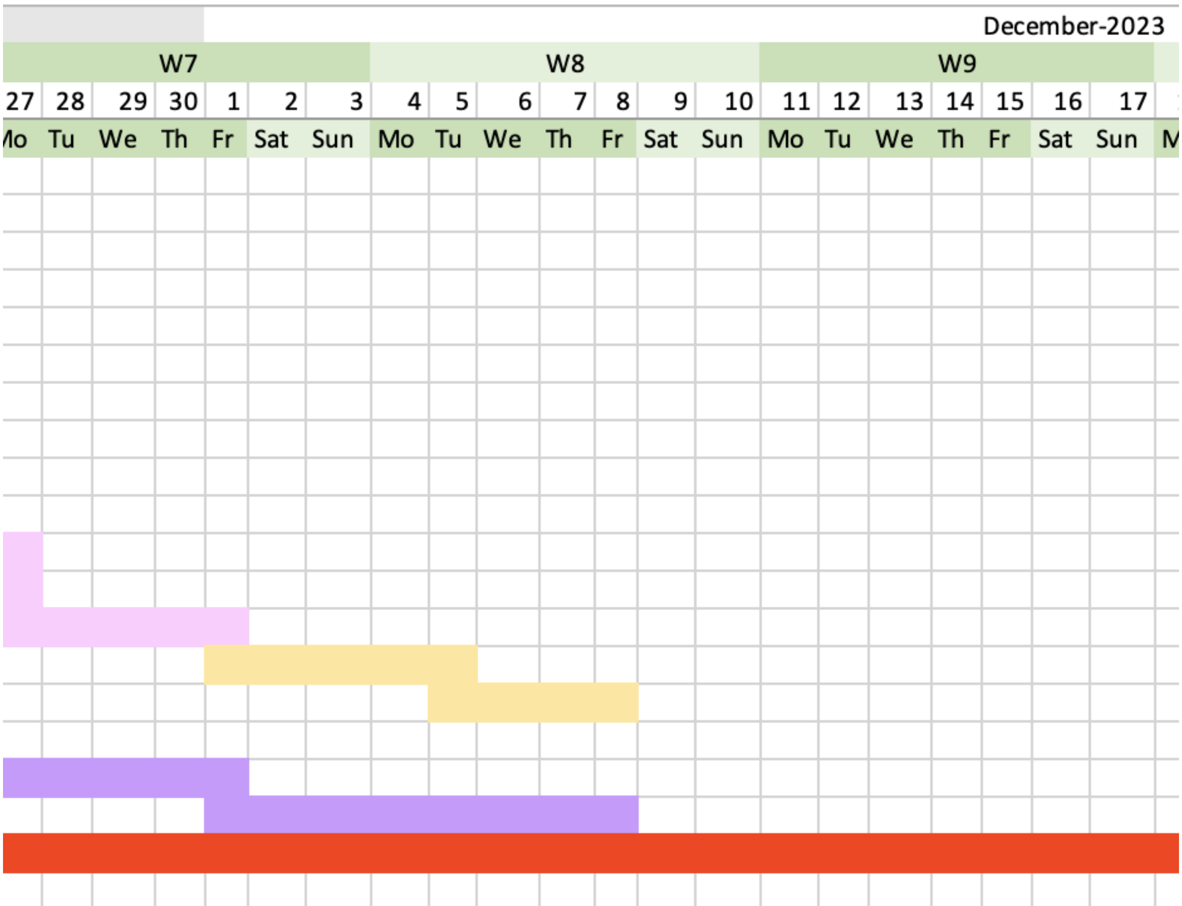
1. Gao, J. (2014), “Machine learning applications for data center optimization”, Google.
2. Lazic, N.; Lu, T.; Boutilier, C.; Ryu, M.; Wong, E.; Roy, B. and Imwalle, G. (2018) “Data center cooling using model-predictive control”, Proceedings of the 32nd International Conference on Neural Information Processing Systems, 3818-3827.
3. Le, V. D. (2021). "Air free-cooled tropical data center: design, evaluation, and learned lessons". IEEE Transactions on Sustainable Computing.
4. Shoukourian, H.; Wilde, T.; Labrenz, D. and Bode, A. (2017), “Using machine learning for data center cooling infrastructure efficiency prediction”, 2017 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), 954-961.
5. Stamatescu, I.; Ploix, S.; Fagarasan, I.; and Stamatescu, G. (2018). “Data center server energy consumption optimization algorithm”. 2018 26th Mediterranean Conference on Control and Automation (MED).
6. Yang, Z.; Du, J.; Lin, Y.; Du, Z.; Xia, L.; Zhao, Q.; and Guan, X. (2021). “Increasing the energy efficiency of a data center based on machine learning”. Journal of Industrial Ecology, 26(1), 323–335.

Section 6 Appendix

	WP	Task Name	Start	Finish	Duration
WP1	T1.1	Literature Review	16.10.2023	23.10.2023	7
	T1.2	Observing KoçSistem Data Center	23.10.2023	24.10.2023	1
	T1.3	Deciding Project Objectives, Scopes, and Methodology	24.10.2023	27.10.2023	3
WP2	T2.1	Data Collection	27.10.2023	30.10.2023	3
	T2.2	Data Preprocessing	30.10.2023	2.11.2023	3
	T2.3	Understanding the Current Features	2.11.2023	3.11.2023	1
WP3	T3.1	Feature Engineering	3.11.2023	8.11.2023	5
	T3.2	Data Analysis	8.11.2023	10.11.2023	2
	T3.3	EDA and Visualization	10.11.2023	17.11.2023	7
WP4	T4.1	Design of The Models	11.11.2023	16.11.2023	5
	T4.2	Developing the Predictive Model	16.11.2023	27.11.2023	11
	T4.3	Training and Testing	16.11.2023	27.11.2023	11
	T4.4	Model Selection	27.11.2023	1.12.2023	4
WP5	T5.1	Sensitivity Analysis of Decision Variables	1.12.2023	5.12.2023	4
	T5.2	Risk Analysis for The Set Parameters	5.12.2023	8.12.2023	3
WP6	T6.1	Designing the Mathematical Model	10.11.2023	24.11.2023	14
	T6.2	Developing the Model	24.11.2023	1.12.2023	7
	T6.3	Results and Decision Making	1.12.2023	8.12.2023	7
WP7	T7.1	Drafting the reports	30.10.2023	19.01.2024	81
	T7.2	Review and Final Report	13.01.2024	19.01.2024	6




November-2023
W4
W5
W6
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
Mo Tu We Th Fr Sat Sun Mo Tu We Th Fr Sat Sun Mo Tu We Th Fr Sat Sun


December-2023
W7
W8
W9
27 28 29 30 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Tu We Th Fr Sat Sun Mo Tu We Th Fr Sat Sun Mo Tu We Th Fr Sat Sun