

**REPUBLIC OF TURKEY**

**AKDENİZ UNIVERSITY**



**ROOM-BASED ENERGY CONSUMPTION PREDICTION WITH MACHINE  
LEARNING IN INDUSTRIAL COLD STORAGE**

**Burak ERKAN**

**INSTITUTE OF NATURAL AND APPLIED SCIENCES**

**DEPARTMENT OF COMPUTER ENGINEERING**

**MASTER THESIS**

**JULY 2021**

**ANTALYA**

**REPUBLIC OF TURKEY**

**AKDENİZ UNIVERSITY**



**ROOM-BASED ENERGY CONSUMPTION PREDICTION WITH MACHINE  
LEARNING IN INDUSTRIAL COLD STORAGE**

**Burak ERKAN**

**INSTITUTE OF NATURAL AND APPLIED SCIENCES**

**DEPARTMENT OF COMPUTER ENGINEERING**

**MASTER THESIS**

**JULY 2021**

**ANTALYA**

**REPUBLIC OF TURKEY**  
**AKDENİZ UNIVERSITY**  
**INSTITUTE OF NATURAL AND APPLIED SCIENCES**

**ROOM-BASED ENERGY CONSUMPTION PREDICTION WITH MACHINE  
LEARNING IN INDUSTRIAL COLD STORAGE**

**Burak ERKAN**

**DEPARTMENT OF COMPUTER ENGINEERING**

**MASTER THESIS**

This thesis unanimously accepted by the jury on 01/07/2021.

Prof. Dr. Melih GÜNAY (Supervisor)

Asst. Prof. Dr. Özge ÖZTİMUR KARADAĞ

Asst. Prof. Dr. Taner DANIŞMAN

Handwritten signatures in blue ink, corresponding to the names listed above. The signatures are stylized and written in a cursive style.

## ÖZET

### ENDÜSTRİYEL SOĞUK HAVA DEPOLAMADA MAKİNE ÖĞRENİMİ İLE ODA BAZLI ENERJİ TÜKETİMİ TAHMİNLEME

**Burak ERKAN**

**Yüksek Lisans Tezi, Bilgisayar Mühendisliği Anabilim Dalı**

**Danışman: Prof. Dr. Melih GÜNAY**

**Temmuz 2021; 36 sayfa**

Bu tez çalışmasında endüstriyel soğuk hava depolarında oda veya depo bazlı enerji tüketiminin tahminlenebilmesi için gereken teorik ve pratik çalışmalar yer almaktadır. Tez çalışmaları aynı zamanda Türkiye Cumhuriyeti Sanayi ve Teknoloji Bakanlığı tarafından desteklenen bir ArGe Merkezi proje çalışmalarının bir kısmını kapsamaktadır.

Merkezi sistem soğutma ile soğutulan endüstriyel soğuk hava depolarında genellikle oda veya depo bazında enerji ölçümü yapılmamaktadır. Oda veya depo bazlı enerji tüketim verileri olmaması operasyonel pek çok çalışmanın verimsiz şekilde yapılmasına sebep olmaktadır. Oda veya depo tüketimleri elde edilerek operasyonel işlerin veriye dayalı olacak şekilde yapılması mümkün hale gelebilir. Bu durumda tez çalışmasında detaylarıyla bahsedildiği üzere büyük oranda enerji tasarrufları yapılabilir.

Bu kapsamda başlatılan ArGe Merkezi projesinin analiz çalışmaları gerçekleştirilmiştir. Analiz çalışmasının sonucunda merkezi sistem ile soğutulan bir tesisin oda veya depo tüketimlerinin elde edilmesi için en büyük bilinmezliğin evaporatör kapasiteleri olduğu anlaşılmıştır. Tez çalışmaları kapsamında makine öğrenmesi teknolojisi derinlemesine araştırılmış ve Feature Importance Scoring yöntemleri ile tesislerde bulunan oda veya depoların soğutucu komponenti olan evaporatör ünitelerinin kapasiteleri farklı algoritmalar kullanılarak tahminlenmiştir.

Tezin yazılım geliştirme çalışmaları sonucunda kapasite tahminlemesi yapan yazılım modülü elde edilmiştir. Bu modül REST Api teknolojisi ile bir web servis hizmeti verecek şekilde tasarlanmıştır. Bu modülün ArGe Merkezi projesinde nasıl bir konumda yer alacağı da tez çalışmasında tartışma ve bulgular kısmında bir diagram olarak yer almaktadır.

**ANAHTAR KELİMELEER: Enerji Tüketimi Tahminlemesi, Soğuk Zincir, Öznitelik Çıkarımı, Öznitelik Skorlama**

**JÜRİ:** Prof. Dr. Melih GÜNAY

Dr. Öğr. Üyesi Özge ÖZTİMUR KARADAĞ

Dr. Öğr. Üyesi Taner DANIŞMAN

## **ABSTRACT**

# **ROOM-BASED ENERGY CONSUMPTION PREDICTION WITH MACHINE LEARNING IN INDUSTRIAL COLD STORAGE**

**Burak ERKAN**

**MSc Thesis in Computer Engineering**

**Supervisor: Prof. Dr. Melih GÜNAY**

**JULY 2021; 36 pages**

In this thesis, there are theoretical and practical studies required for room and/or warehouse-based energy consumption prediction studies in industrial cold storage. Thesis studies also cover a part of the Research and Development Center project studies supported by the Republic of Turkey Ministry of Industry and Technology.

In industrial cold storages cooled by central system cooling, energy measurement is generally not made on a room and/or warehouse basis. The absence of room and/or warehouse-based energy consumption data causes many operational works to be done inefficiently. By obtaining room and/or warehouse consumption, it becomes possible to carry out operational works based on data. In this case, as mentioned in detail in the thesis study, a great deal of energy savings can be made.

Analysis studies of the Research and Development Center project initiated in this context were carried out. As a result of the analysis study, it has been understood that the biggest unknown in obtaining the room and/or warehouse consumption of a facility cooled by the central system is the evaporator capacities. Within the scope of the thesis studies, machine learning technology has been researched in depth. With Feature Importance Scoring methods, the capacities of the evaporator units, which are the cooling components of the rooms and/or warehouses in the facilities, were predicted using different algorithms.

As a result of the software development studies of the thesis, a software module that makes capacity estimation has been obtained. This module is designed to provide a web service with REST API technology. The position of this module in the Research and Development Center project is also included as a diagram in the discussion and findings section of the thesis study.

**KEYWORDS: Energy Consumption Prediction, Cold Chain, Feature Selection, Feature Importance Scoring**

**COMMITTEE: Prof. Dr. Melih GÜNAY**

Asst. Prof. Dr. Özge ÖZTİMUR KARADAĞ

Asst. Prof. Dr. Taner DANIŞMAN

## ACKNOWLEDGEMENTS

I wish to thank my committee members for their time and dedication. A special thanks to Prof. Dr. Melih GÜNAY my chair and advisor, for believing in me and taking me on ,and my company Cantek Group, who showed every kind of understanding and support during my graduate education and thesis work.

I would like to thank my dear friend Mehmet Arif Emre ŞEN, who provided support and assistance throughout my work and helped me with my design work day and night, Emre ALTUNTAŞ who helped me with my translation work, and my colleagues Fırat ATAŞ and Nuri ÖZÇELİK who were always there for me with their support.

I would also like to thank Gülcihan ERKAN, my mother, who did not spare her financial and moral support for us to come to this day, Tekin ERKAN, my father, who devoted his life to his children, Gökhan ERKAN, my older brother, who always supported us with his presence,

I would like to express my endless thanks to Ayla ERKAN, my beloved wife, who showed her understanding and support throughout my thesis as she was always there for me under all circumstances, and our dear daughter Eylül ERKAN for her understanding during my work.



## LIST OF CONTENTS

ÖZET . . . . .	i
ABSTRACT . . . . .	iii
ACKNOWLEDGEMENTS . . . . .	v
TEXT OF OATH . . . . .	vii
LIST OF ABBREVIATIONS . . . . .	viii
LIST OF FIGURES . . . . .	ix
LIST OF TABLES . . . . .	x
1. INTRODUCTION . . . . .	1
2. LITERATURE REVIEW . . . . .	3
2.1. The Importance Of Cold Storage In Human Life . . . . .	3
2.2. The Energy Methods Used in Cold Storage . . . . .	4
2.3. Cold Storage Capacities of Countries . . . . .	4
2.4. Differences in Energy Consumption of Cold Stores and Their Reasons . . . . .	5
2.5. Research on Artificial Intelligence Studies on Energy Consumption Predict- ing in Air Conditioning or Similar industries . . . . .	7
2.6. Artificial Intelligence . . . . .	10
2.6.1. Chronological History Of Artificial Intelligence . . . . .	11
2.6.2. Overview Of Artificial Intelligence And Machine Learning . . . . .	12
2.6.3. Introduction To Artificial Neural Networks . . . . .	13
2.6.4. Artificial Neural Networks Structural Origin And Basic Elements . . . . .	15
2.6.5. Studies On Energy Forecasting In Buildings . . . . .	16
3. MATERIAL AND METHOD . . . . .	20
3.1. Gathering Data . . . . .	20
3.1.1. Data preparation . . . . .	21
3.1.2. Analyzing Data . . . . .	22
3.1.3. Data Wrangling . . . . .	23
3.2. Train The Model . . . . .	24
3.3. Test the model . . . . .	25
3.4. Deployment . . . . .	25
4. RESULTS AND DISCUSSION . . . . .	27
4.1. Results . . . . .	28
5. CONCLUSION . . . . .	32
6. REFERENCES . . . . .	33
CURRICULUM VITAE	

### **Text Of Oath**

I declare that this study "On the use of quantum cryptography in financial it systems", which I present as master thesis, is in accordance with the academic rules and ethical conduct. I also declare that I cited and referenced all material and results that are not original to this work.

01/07/2021

Burak ERKAN



## **LIST OF ABBREVIATIONS**

AI	: Artificial Intelligence
ANN	: Artificial Neural Network
BTU	: British Thermal Unit
MAE	: Mean Absolute Error
MAPE	: Mean Absolute Percentage Error
MSE	: Mean Squared Error
REST	: Representational state transfer

## LIST OF FIGURES

<b>Figure 2.1.</b>	Energy Saving Potential . . . . .	6
<b>Figure 2.2.</b>	Photovoltaic Panels . . . . .	9
<b>Figure 3.3.</b>	Prediction Input Output Diagram . . . . .	24
<b>Figure 3.4.</b>	Component Diagram of Room-based Consumption Prediction . .	26
<b>Figure 4.5.</b>	Pilot Facility Plan . . . . .	27
<b>Figure 4.6.</b>	Microservices Flow Diagram . . . . .	28
<b>Figure 4.7.</b>	Feature Importance Scores - Result of RESTful API . . . . .	29
<b>Figure 4.8.</b>	Facility Technical Details . . . . .	29
<b>Figure 4.9.</b>	Regression - Feature Importance Scores - Evaporation Weights .	30

## LIST OF TABLES

<b>Table 2.1.</b>	Comparison of Methods . . . . .	9
<b>Table 2.2.</b>	History of Artificial Intelligence . . . . .	12
<b>Table 2.3.</b>	Comparison of Traditional Algorithms and ANNs . . . . .	14
<b>Table 2.4.</b>	Biological Neural Structure and ANN Correspondence . . . . .	16
<b>Table 2.5.</b>	The comparison table of ANN and nonlinear regression model . . . . .	18
<b>Table 4.6.</b>	Ratio Prediction Results for Linear Regression . . . . .	31

## 1. INTRODUCTION

Industrial cooling systems are generally designed in two types as split and central system structures.

In the split system setup, the evaporator and condenser units are separated from each other, and the condenser unit keeps its compressor and provides the cold fluid you need for cooling of the evaporator unit, according to your needs. The condenser and evaporator, which are split system components, are controlled by an autonomous controller. To give an example of the split cooling system, it can be thought that it is designed like the split air conditioners that are standard in most houses. In the central system setup, there is usually more than one cold room or warehouse, and the cooling of these cold rooms or warehouses is provided by at least one Evaporator unit in them.

The central cooling unit usually has more than one compressor and it supplies the required cold fluid by operating these compressors gradually in line with the need from the evaporators. Technically, the heat drawn from the evaporator(s) is thrown to the outside through the condenser(s). The central cooling unit has no relation with the evaporators in the rooms except the pressure relation, it is managed by an autonomous controller in the components board within the central cooling unit. To give an example of the central cooling system, it can be thought that it is similar to the multi-split air conditioning system, which is increasingly used in homes.

In the prediction study, the power meter and central cooling unit controller data on the board of the central cooling unit and the data of the room controllers in the facility are used as learning data efforts are made to predict the cooling capacities and consumption of rooms or warehouses without the need for the use of energy measuring devices in rooms or warehouses. The biggest uncertainty for any work to be done in this context is the unknown of the capacity of the evaporators, one of the cooling equipment of the room or warehouse. Although this information is technically available, keeping this technical information in industrial control systems and the estimation studies that can be made on this information are called rule-based software. The subject of the thesis is to determine the cooling capacity required to make room or warehouse-based energy consumption estimation with machine learning. The theoretical knowledge and gains to be obtained from the thesis work will be used in the project carried out within the scope of Cantek Group

Research and Development Center and will be used in the room or warehouse energy consumption estimation project.

In the prediction study, the power meter and central cooling unit controller data on the panel of the central cooling unit and the data of the room controllers in the facility are used as learning data. Efforts are made to predict the cooling capacities and consumption of rooms or warehouses without the need for the use of energy measuring devices in rooms or warehouses.

## **2. LITERATURE REVIEW**

### **2.1. The Importance Of Cold Storage In Human Life**

When the history of humanity is examined, it is understood that one of the most important criteria of being a developed society is food production and supply. Developed countries have been well organized in agriculture and animal husbandry throughout history and have produced or supplied the food they need. (Eroğlu et al. 1994)

The ability of countries to be self-sufficient in terms of food has always been an important factor, but a more important competence has been added than being able to produce in the 20th and 21st centuries. This is to meet the food needed with the least participation of manpower. It has become an important success criterion in terms of the development criteria of the countries where the percentage of manpower is used to produce the total food needed by the countries. The less labor is used in the total food production needed, the more industrialized and successful the countries are.

It can be said that about 5% or less of the population of countries that have successfully realized the fourth stage of agricultural transformation work for food production or supply. (Peter Timmer 1988; Arzova et al. 2014)

Today, the most important factor seen as the international level of development in the food and storage sector is no longer production. It is important to transfer agricultural products most efficiently after harvest and to store them in inappropriate conditions.

It is seen as a development in the field of food to make the products produced not only in the season but also in other times of the year. It is necessary to grow agricultural products using high technology, with low labor force participation, and to store them in the best conditions by using high technology. The biggest indicator of development in the field of food is to grow food in the best conditions and store it in the best conditions, to provide a continuous supply of food products, and to create price stability. (Utku 2015)

Being able to store the produced foods well has not only enabled the efficient consumption of the produced food products but also created the opportunity to obtain maximum economic benefit from the produced foods. If the products produced are not stored efficiently, they become garbage from the moment they expire. It is estimated that if the storage conditions of the food products produced in the world are corrected, the wasted food will be enough for the entire African continent. (Utku 2015) This gives us an idea



of how important it is to store food products under appropriate conditions.

What end-users in the food ecosystem demand is that products are always of high quality. In this regard, cold storage is of great importance. (Wu et al. 2018)

## **2.2. The Energy Methods Used in Cold Storage**

The method generally used in cold storage is central cooling units with electric compressor(s). With this method, the needs of the evaporators in the warehouses are met by sending the cold fluid to the cold stores according to the needs. Through evaporators, heat is drawn from the warehouses and released to the outside, thus cooling is provided. All equipment used in this cooling method uses electrical energy.

Cooling systems contain differences. There are also different types of cooling methods from the type of cooling described in this section. The subject of this thesis is the prediction of the energy used in cold storage on a warehouse or cold room basis. For this reason, it is sufficient to transfer cooling systems that use electrical energy, which is the most used method among cold storage types.

## **2.3. Cold Storage Capacities of Countries**

It is known that there are approximately 17 million refrigerated air warehouses in the European continent and the total capacity of these warehouses is approximately 65 million m<sup>3</sup>. It is known that 67% of this capacity is small-scale warehouses of less than 400 m<sup>3</sup>. This shows us that a very large part of the total capacity is formed by warehouses larger than 400 m<sup>3</sup>.

According to the determination made by the Energy Technology Savings Unit (ETSU) in the UK, cold stores consume an average of 34 kWh of energy on a cubic meter basis. (The UK - ETSU 1994)

This amount obtained as a result of the research shows that it has the potential to increase the consumption amount up to 124 kWh if regular controls and improvements are not made. From the example given here, it is understood how important follow-up and optimization are in cold storage. It is known that systems that are followed, controlled, maintained, and optimized by competent people perform well. In addition, maintenance-free systems can consume almost four times more energy compared to optimum operating

conditions. This ratio is so great that a sloppy country in cold storage can waste enough energy for four countries close to its capacity. Everyone knows how precious energy is today, the damage done to the world while producing energy. Environmentally friendly and renewable energy sources are not yet widespread in the world. In this case, the most effective and responsible behavior should be to use the energy produced more effectively.

#### **2.4. Differences in Energy Consumption of Cold Stores and Their Reasons**

The reasons can be the cooling system, defective components, etc. In central cooling systems, which is the method generally used in cold storage, the energy consumption is mostly realized by the central cooling unit. Large-capacity compressors used in the central cooling unit require high energy consumption.

Evans and others (2014)'s study showed that as energy consumption increases in storage facilities, the size of the savings that can be achieved can increase in direct proportion. They analyzed that there is a possibility to save 15% to 40% of energy by supplementing the cooling systems in the facilities with energy-efficient equipment or by improving their configuration. It has been understood that the amount of savings that can be made in some special cases reaches 82%.

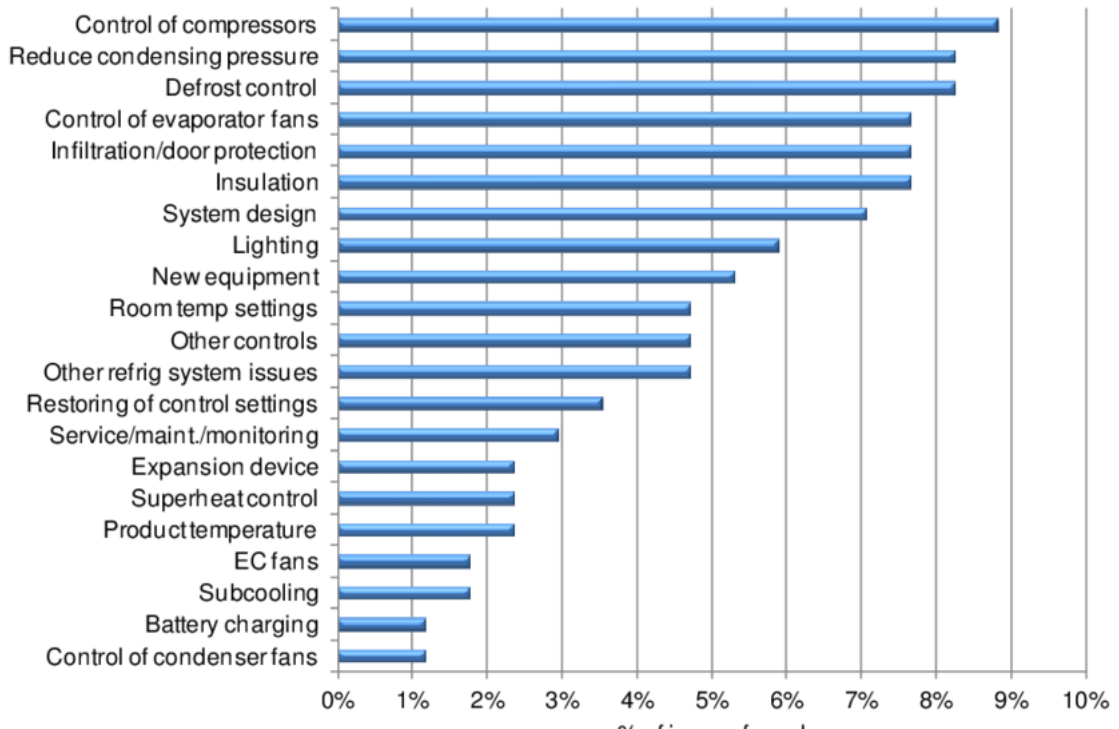
Wu et. al (2018)' analysis showed that energy savings can be made between 8% and 28% in storage facilities with a capacity of more than 100 m<sup>3</sup> and between 8% and 72% in storage facilities with a capacity of more than 100 m<sup>3</sup>.

The energy wastage resulting from the failure of critical components in large cold storage facilities is so great that often the expense of repairing a system pays off within a year at most. (Evans et al. 2014) From this example, it is possible to understand how large the wasted energy can be.

Likewise, from the studies conducted by Werner and others, it is understood that as the cooling capacity grows and the energy consumed increases, the potential for energy saving increases at the same rate. (Werner et al. 2006)

Figure 2.1. showing the ratio of energy savings to total energy consumption that can be achieved through improvement and maintenance work in cold stores cooled by central cooling;

An important point to note here is that the effect of the capacity of the cooling facilities on the savings that can be made can be very different. It is understood that the size of



**Figure 2.1.** Energy Saving Potential

the cooling facilities greatly affects the amount of savings that can be made in which equipment is improved or in which settings optimization is made.

In the study conducted by Evans and Giegel, the percentage analysis of energy savings that can be achieved by the application or improvement of twenty-one cooling components or different methods in storage systems is shown. The savings that can be made here are divided into two as below and above 100 m<sup>3</sup> in terms of capacity. Because, as mentioned before, the gain that can be obtained from large cooling systems in terms of capacity is realized at much greater rates.

As it is easily understood from the analysis made, the greatest energy saving is possible by making good service maintenance and monitoring services and making sure that the system works in optimum conditions.

With good service, maintenance, and monitoring services, savings of around 15% can be made in storage facilities larger than 100 m<sup>3</sup>.

The savings rate that can be made from place to place can reach up to 35%. As a result of the research, it has been recorded that savings of around 5% can be made when the same services are properly performed in facilities smaller than 100 m<sup>3</sup>.

At this point, the importance of monitoring systems and the importance of analyzing the distribution of cooling over the rooms in the facilities cooled by the central cooling system emerge. As it can be understood from the literature study, the most successful optimization method and the method in which the most savings can be made is the analysis-based method. As Evans and Gigiel have determined, only domain experts can detect the problems experienced in these systems accurately and quickly. To make these determinations, these people also need monitoring systems. It is an important condition that cold stores can analyze the working conditions, working times, and environment variables of the cooling systems through these systems. (Evans and Gigiel 2007)

## **2.5. Research on Artificial Intelligence Studies on Energy Consumption Predicting in Air Conditioning or Similar industries**

When the literature on energy prediction is examined, power prediction studies using artificial intelligence algorithms have been encountered.

In her study, İzmirli Ayan (2018) made power predictions on photovoltaic solar panels by using an artificial neural networks algorithm, one of the artificial intelligence methods. With his research, he revealed the change of power produced by a sample photovoltaic panel according to environmental factors. It is understood that ANN uses The Multiple Regression Analysis methods to compare its performance. In this study, solar radiation (W/m<sup>2</sup>), wind speed (m/s), wind chill (°C), humidity(% RH), ambient temperature (°C), and panel temperature (°C) values as artificial neural network data has been collected. The prediction success of the multilayer and feed-forward back propagation artificial neural network and other algorithms used in the artificial neural network created by analysis and experiments has been tested.

İzmirli Ayan (2018), who has been working on the conversion of solar rays, which is a renewable and clean energy source, into electrical energy using photovoltaic solar panels, was able to estimate the output power with a success rate of 98% in his estimations using Artificial Neural Networks. For the power prediction of the photovoltaic system, the prediction capabilities of multiple regression and artificial neural networks (ANN) were compared. It was understood that the artificial neural network model explained the power output with 98% success and the multiple linear regression model 94.8%. As a result of the study, it has been understood that artificial neural networks can be a more successful

method than multiple regression analysis for power estimation of photovoltaic systems. (İzmirli Ayan 2018)

In their research article, Dandıl and Gürgen (2019) carried out the study of estimating the power output of photovoltaic panels and comparing them with heuristic algorithms. In the study, photovoltaic panels were positioned at different angles of 10°, 20°, 30°, 40°, 50°, 60°. They took the average of the ten-minute values every ten minutes over the panels positioned at different angles. They used the data collected in ten-minute averages as a training set. The collected data was then used as a training dataset on the model. Monthly energy consumption estimation was made with the model trained on this dataset. During the estimation studies, Artificial Neural Networks (ANN), which can learn with different algorithms, was used for the monthly estimation of the power values obtained from the PV panels. Unlike the estimation methods known in the literature, it has been understood that ANN models are based on Particle Swarm Optimization (PSO) optimization algorithm and ANN models trained using the Backward Elimination method and Clonal Selection Algorithm (KSA) are utilized. As a result of the study, it was concluded that the ANN structure trained with PSO was more successful than the ANN structures trained with KSA and GR algorithms. It was understood that the effectiveness of the methods used on the results obtained was confirmed by the analysis of the average percent error between the actual values measured and the estimated values. (Dandıl and Gürgen 2019)

For electricity consumption in our country, end users have also been allowed to become free consumers. Subscribers who use a certain amount of electricity above the monthly average consumption can become customers of different electricity suppliers by submitting a petition.

Electricity institutions responsible for distribution regions are obliged to submit the end-of-month indexes of their eligible consumers to electricity suppliers. The end-of-month index information of all eligible consumers is not known by the electricity authority. In the study of Özlü (2018), the end-of-month index values of the relevant eligible consumers were obtained by estimation method using artificial intelligence algorithms. The obtained values were compared with the index estimates obtained by applying the consumption estimation methodology determined by the Turkish Energy Market Regulatory Authority and the success of the artificial intelligence methods used was compared. In the study, it was understood that the artificial bee colony algorithm (ABC) optimization



**Figure 2.2.** Photovoltaic Panels

method was used for time series prediction in artificial neural networks. In her study, Özlü (2018) used artificial intelligence algorithms to solve the problem faced by electricity distribution companies while estimating by using a real artificial bee colony algorithm. A significant improvement has been achieved with the new estimation method.

As a result of the studies, it has been understood that an approach based on the Artificial Neural Network method can be used as a method of estimating the end-of-month values of electricity meters, and more successful estimations can be obtained.

**Table 2.1.** Comparison of Methods

Average Performance (%)	ABC-ANN	Current Method	Difference
House	93.4	79.3	14.1
Business	86.3	65.4	20.9
Industry Establishment	97.2	81.1	16.1

With an approach based on the Artificial Neural Network method, it has been understood that it has been concluded that it makes 14.1% more successful estimations in residences, 20.9% in commercial establishments, and 16.1% in industrial establishments

compared to the currently used method. (Özlu 2018)

Dong and others (2021), with the hourly energy consumption estimation they made in an office building in the USA - New York, aimed to estimate with a higher level of consistency than the studies that can be done with a normal artificial intelligence algorithm. To achieve this, they used ensemble learning and energy consumption pattern classification, and as a result, they achieved a more successful estimation result than traditional methods. (Dong et al. 2021)

## **2.6. Artificial Intelligence**

Artificial Intelligence (AI) studies were first made in 1943 by McCulloch and Pitts. Artificial Intelligence software imitates the learning process by imitating people's thinking and evaluating the results obtained from past experiences. Since people act through the learning process, more successful results can be obtained in solving problems compared to rule-based software. (Erler et al. 2003)

The term "artificial intelligence" was first coined at a conference at Dartmouth College in Hanover, New Hampshire, in 1956. (Haenlein and Kaplan 2019)

With the development of technology, the processing speed of computers has increased. This capability has now enabled computer software to perform more operations in a shorter time. Starting from this point, it has become possible for computer software to collect information, make inferences and make decisions by imitating the working structure of the human brain. In some cases, computer software can try to find solutions using heuristic methods, even in cases that cannot be mathematically put into the equation. Studies that equip computer software with such capabilities, produce solutions to problems by using existing abilities, and develop original methods in this direction are generally known as "artificial intelligence" studies. (Elmas 2007)

The development of Artificial Intelligence has constantly changed and developed after it was put forward as an idea in the 1950s. One of the factors that accelerate this change is the use of Artificial Neural Networks. Especially after the 1990s, there have been great developments in artificial neural networks. This technology attracted the attention of scientists in those years and has managed to be one of the most attractive subjects for scientists since those years. This interest has led to the transformation of this technology into one of the most interesting branches of science. Studies on artificial neural networks

and artificial intelligence have become a part of real life, rather than just studies by scientists in laboratories. Computer software using artificial intelligence algorithms is used for the solution of many complex problems experienced in daily life. One of the biggest triggers of this rapid transformation in technology is the effect of the internet on the world of technology and humanity. The Internet has accelerated access to information at an incredible rate, paving the way for us to go beyond being an industrial society and become an information society. (Öztemel 2012)

Artificial neural networks have attracted the attention of scientists because they work in a structure similar to the human brain's learning and learning way, creating and generating new information and discovering new information in this way. By working in the field of artificial neural networks of artificial intelligence, scientists have applied this technology to different fields and used it in solving complex problems. At the same time, they used Artificial Neural Networks as an indirect way to understand the human brain, which is full of uncertainties. As it is known, there are still hundreds of studies on the human brain and how it works, consciousness, and learning. These issues continue to be among the greatest unknowns of humanity.

It is known that Albert Einstein reached mathematical conclusions about the existence of black holes during his work on matter, time, gravity, and space in his article published in 1939. In his work spanning years, he stated that this phenomenon cannot exist in real life. Einstein (1939) Many years after Einstein's this determination, in 2012, scientists found evidence of the existence of black holes. As in this example, science may find the answer to the questions of how the human brain works in the future, what is real consciousness and how it is formed through Artificial Intelligence studies. (Orosz et al. 2011)

### **2.6.1. Chronological History Of Artificial Intelligence**

History of Artificial Intelligence;



**Table 2.2.** History of Artificial Intelligence

Year	Work Done
1943	McCulloch and Pitts: Boolean circuit model of the brain
1950	Turing's "Computer-processing machines and intelligence"
1956	Dartmouth Interview: The name "artificial intelligence" was coined
1952-1969	IBM wrote the first program that could play chess. The first international conference on artificial intelligence was held
1950-1959	Newell and Simon's logic theorist, Gelernter's geometry engine
1965	A complete algorithm developed by Robinson for logical thinking
1966-1973	Artificial intelligence encounters computational complexity. Neural networks research is pretty much lost
1969-1979	The first development steps of knowledge-based systems
1980	Artificial intelligence becomes an industry
1986	Neural networks have become popular again
1987	Artificial intelligence becomes science
1995	Clever agents (to use the term) emerge
1997	Deep Blue defeated Kasparov
1998	With the widespread use of the Internet, many artificial intelligence-based programs have reached large audiences
2000-2005	Robot toys were launched

### 2.6.2. Overview Of Artificial Intelligence And Machine Learning

Artificial Neural Networks (ANN) are a sub-research field of Artificial Intelligence in general. It covers the studies done to learn computer and computer software. Systems containing Artificial Intelligence (AI), in which artificial neural networks are used, are used in many areas of daily life, such as engineering automation systems, medicine, cybersecurity, etc. The usage areas of solutions containing Artificial Intelligence are expanding day by day.

The categorical classification of Artificial Intelligence systems is as follows;

- Expert Systems
  - Expert systems are systems that generally provide solutions to problems by making inferences using relationships between different features or information.
- Artificial Neural Networks (ANN)
  - They are systems that learn on the inputs and results of existing examples, and then take the information of events that have never happened as input and can predict results with inferences obtained from examples that have been experienced before.

- Genetic Algorithms
  - It is a method of using different solutions together for problems that cannot be solved by applying traditional optimization methods.
- Fuzzy Propositional Logic
  - It is a method that facilitates decision-making in situations that cannot be expressed rationally.
- Intelligent Factors
  - These are systems where different Artificial Intelligence methods can be programmed and used in different combinations.

Machine Learning (ML) in general, is a set of systems that help to improve the results by the method of the learning process through the inputs and results of past trials and changing behaviors during new trials. The learning process mentioned here takes place in different ways with the use of different strategies.

Learning strategies from past essays are divided into the following headings;

- Supervised Learning
- Assisted Learning
- Unsupervised Learning

### **2.6.3. Introduction To Artificial Neural Networks**

Artificial Neural Networks(ANN) are computer systems that mimic the learning function of the human brain. They perform the learning process through examples, and each learning creates a weight value in the connections between the Artificial Neural Networks. The knowledge learned is stored in this value.

The biggest innovation achieved is that the method used to calculate this value is resistant to situations such as incomplete information or uncertainty. It's just an algorithm, that is, the learning process can adapt to any situation and condition. To give an example for this, an adult is likely to know what it means when they see a stop sign. The presence of a tree trunk in front of this sign often does not change the result. People can identify

this sign even if the image is missing because they have experienced this information hundred of times in the past, and even if there are missing parts, the brain can predict the whole. The same is true in neural networks, when the algorithm learns well enough, it can complete the missing pieces and overcome the uncertainties. This can be considered as proof that a computer or software, which is a machine, successfully imitates the learning ability of the human brain.

Artificial Neural Networks are widely used in classification, pattern recognition, signal filtering, data compression, and optimization studies. Because there are examples of high success using Artificial Neural Networks in these areas in the past. Artificial Neural Networks technology is actively used in daily life in many areas such as determining the most appropriate route with navigation, image recognition, and medical analysis devices.

All kinds of developments in the field of artificial intelligence and hardware pave the way for the use of more artificial intelligence technology in daily life. This technology has brought a different perspective and solution method to problems that were seen as unsolvable in the past.

**Table 2.3.** Comparison of Traditional Algorithms and ANNs

<b>Traditional Algorithms</b>	<b>Artificial Neural Networks</b>
Outputs are obtained by applying the rules set in their inputs. (rule-based)	During learning, input and output information is given and rules are set.
Information and algorithms are precise.	It benefits from the experience.
Calculation; central, synchronous and sequential.	Calculation; batch, asynchronous and parallel after learning.
Memory is packed and literal stored.	The memory is allocated and spread over the network.
There is no fault tolerance.	It has fault tolerance.
It is relatively fast.	It is slow and hardware-dependent.

(Pirim 2006)

#### **2.6.4. Artificial Neural Networks Structural Origin And Basic Elements**

As the name suggests, Artificial Neural Networks are designed on a structural basis, inspired by the biological nervous system. The biological nervous system is like a network that is constantly communicating with each other. These nerve cells use connection pathways called synapses and axons for intra-network communication. They send information to other nerve cells via the axon and information from other nerve cells is collected via dendrites. Synapses are like connecting units for communication between axons and dendrites.

Artificial Nerve Cells also have a data acquisition function, they receive data with it and share the result obtained by passing it through the activation function with other cells.

Artificial nerve cells can be roughly said to have three layers. These layers are listed below;

- Input Layer
- Middle Layers
- Output Layer

The input layer is the entry point of the information in the neural network. The information coming to the network is processed in the middle layers, that is, it becomes a result of interacting with the weight values existing in the network. As a result of this process, the weight values existing between the artificial nerve cells change, and the Artificial Neural Network (ANN) learning process is realized. The weights that change and develop between the artificial nerve cells as a result of the process that takes place here represent the learning process of the network.

The common method used when training the Neural Network is to randomly allocate some of the samples for the learning process. The remaining examples are intended to be used for testing purposes after the learning process. In this study, if the Artificial Neural Network gives correct answers to the test samples, the network is considered to have learned. When a Neural Network that learns with training samples is tested with test samples, the learning success of the network is measured rationally. Weights have been formed in the network and learning has taken place, but when these weights are examined, the values will not mean anything singularly. It is not possible to understand what value

**Table 2.4.** Biological Neural Structure and ANN Correspondence

<b>Biological Nervous System</b>	<b>Artificial Nervous System</b>
Neuron	Processing element
Dendrite	Addition function
Cell body	Transfer function
Axons	Artificial neuron output
Synapses	Weights

means or the difference between success or failure by looking at the values. For this reason, artificial neural networks can also be compared to black boxes. The artificial intelligence formed is hidden in the values formed inside this black box.

Learning performance is tried to be increased through different models created by making structural changes in Artificial Neural Networks (ANN). The different models obtained here are called artificial neural network models.

The factors that determine an artificial neural network model, in general, are listed below;

- Neural Network Topology
- Summation Function
- Activation Function
- Learning Rule And Strategy

(Öztemel 2012)

#### **2.6.5. Studies On Energy Forecasting In Buildings**

Research by Zhao and Magoules (2012) has shown that many factors affect the energy performance of structures. Some of these are ambient conditions, air conditioning conditions, and air circulation flow. In addition to these, lighting, HVAC systems can also be added. As stated, there are many factors that affect the energy consumption of a building, in this case, many parameters and environment variables make it difficult to make this estimation while estimating the energy consumption of the building. Zhao and

Magoules (2012) have developed several models to solve this problem in their research. These models include detailed engineering approaches and statistical methods.

As a result, it was emphasized that in the studies conducted for the estimation of the energy consumption of the buildings, estimation was tried to be made in an environment where uncertainty increased due to the involvement of many factors. This complexity and uncertainties made the solution of the problem difficult. For this reason, many solution models have been tried in the past for the solution of this problem. Zhao and Magoules (2012) compared these approaches in their review study.

There are detailed engineering solutions included in the scope of the review. Some of the commonly used methods are; Statistical models, data science studies with artificial intelligence, artificial neural network, and Support Vector Machines. Zhao and Magoules (2012) based their research on using these solution models on new estimation problems. To achieve better performance by using the mentioned models, studies have been carried out to optimize the model parameters or input parameters. Some of the aims of this study are to simplify the problem or develop new models under certain conditions, as well as to compare existing models. As a result of the comparison, it has been understood that each model has its advantages and disadvantages. It is difficult to understand exactly which model is more successful than the others, since a comprehensive comparison cannot be made under the same conditions. As a result of the research, it is understood that extensive studies on artificial intelligence solutions continue. It is stated that there are new, comprehensive, and promising studies covering artificial intelligence algorithms. It is understood that the models developed using data science bring different alternatives to the solution models and they continue to be developed. It has been emphasized that models containing energy estimations of structures in which artificial intelligence technologies are also used are promising. (Zhao and Magoules 2012)

The rapid growth in the human population also causes a rapid increase in electricity consumption. Therefore, efficient energy management and energy consumption estimation are more important than ever. Proactive energy management is required for all studies to be able to save energy in buildings where a large amount of energy is consumed. For energy planning and management, it depends on accurate energy need estimations.

In their research, Ahmad and others (2014) used The Support Vector Machine method and artificial neural network (ANN) methods from artificial intelligence methods to make

eastern energy estimation in buildings. Both methods are widely used for energy consumption estimation in buildings. With the use of artificial intelligence methods, it is aimed to make the closest estimations to the real energy needs, unlike traditional methods. In this study, by using a hybrid method, the accuracy of the predictions has been tried to be increased to use Support Vector Machine and artificial neural network (ANN) methods together. (Ahmad et al. 2014)

**Table 2.5.** The comparison table of ANN and nonlinear regression model

<b>Year</b>	<b>Actual</b>	<b>ANN</b>	<b>Regression</b>
2000	15342.46	15283.92	16792.82
2001	14856.34	14561.08	16188.65
2002	14370.23	14351.00	15583.68
2003	14574.89	14073.28	15094.45
<b>MAPE Error</b>		0.0099	0.075

As can be seen from the table, the error margin of MAPE decreased by approximately 6/7.5 in the estimation made by the ANN method compared to the estimation made by the Regression method.

Prediction models are of great importance for energy control operational processes and energy management of buildings. In the literature, energy estimations of structures are generally divided into three categories;

- white-box (physics-based)
- black-box (data-driven)
- gray-box (combination of physics-based and data-driven)

It is understood that Wu and others (2018) 's articles focused on energy estimation models in the structural field. Examples include short-term climate and weather forecasting. In their studies, there are prediction models on the energy control of single or multiple structures under optimum conditions. They conducted research on the up-to-dateness of these models and published them in summary. In addition, model-based and model-free optimization methods used for energy estimation of structures were evaluated and compared with each other. They also examined the superior forecasting success brought by

the use of agent-based modeling as a new modeling strategy in recent years. (Amasyali and El-Gohary 2018)



### 3. MATERIAL AND METHOD

As mentioned in the literature research, it is known that 60-70% of the energy of a facility is realized by the central cooling unit. (Evans et al. 2014) To estimate this consumption on a room basis, the ignorance of the capacities of the evaporators, which are the cooling components in the rooms where the active cooling is carried out, that is, the heat load from the rooms is collected, has become a big problem. In this study, in order to make room-based energy estimation in a cold storage facility, firstly, the need to obtain the cooling system structure installed in the facility by Feature Selection methods has emerged. With the regression studies, the weights of these evaporators compared to each other were determined. Room-based energy consumption values can be obtained with a rule-based software service using the in-plant evaporator weight values to be obtained.

#### 3.1. Gathering Data

The data of the Cantek Cloud Monitoring System (Octocloud), which was specially developed for Cantek Group, which can be used in the field of industrial cooling in general but wherever similar data types are kept, was used as a source to make energy estimation in industrial storage, which is the targeted output. Octocloud ecosystem is designed and works according to microservice architecture. Data held in a NoSql database solution in the cloud are integer, decimal, and bool. The data source is the controllers that work autonomously and initialization is made operationally. Data from the controllers, which are the data source in the Octocloud ecosystem, are read with an Octolive device that acts as middleware. In the Octocloud ecosystem, there is a microservice called Alfred, where workflows related to incoming data are carried out. In general, the purpose of Alfred microservice is to receive data independent of controller and data, dependent on data type and communication protocol (modbus-rtu, modbus-tcp, canbas, etc.), and store it in NoSql database solution. The types of controllers that can be communicated must have been previously introduced to the system via the interface and updated to the device called Octolive, which acts as a gateway. At this point, the sensor data, which is called raw data, and/or the values kept by the workflow of the autonomous controller, fall into the NoSql database, which is the persistence environment of the cloud monitoring system and is ready to be used in data visualization, analysis and/or data studies.

### 3.1.1. Data preparation

Data visualization or raw use of data required for forecasting in data science studies makes processes long and tedious. It has been observed as a result of the literature research that it can take days to train some machine learning algorithms. [Shaw 2008] This situation causes data science studies on cloud servers, which is a common method for environments where software runs today, to become costly. At this point, it has been tried to be overcome by obtaining summary data (metadata) from raw data with the development of a microservice developed within the Octocloud system that converts the data into summaries such as 5 minutes, 60 minutes, and daily summaries. As a result of this study, it can be summarized as turning raw data into a one-line summary of 5-300 lines of data from the field in 5 minutes. The operation performed and the results obtained vary according to the type of raw data. The average of 5 – 300 data obtained for integer and/or decimal data types, the average of the values taken when the initialization is active, the lowest and highest values recorded within 5 minutes are calculated and stored separately in the system as summary data. The 5-minute summary data obtained as a result of this study is then used as the basis for calculating hourly and daily summaries.

To extract summary data from raw data in the summary microservice in the Octocloud ecosystem, a different method is used to obtain metadata for bool values. The most difficult information to obtain over bool data is time information. It is relatively simple to calculate run times on data from the data source routinely during a given interval. At this point, working times from bool values can be calculated with a simple mathematical operation based on the frequency of data extraction. However, the systems used in real life do not work with a standard data extraction frequency due to different reasons.

Octolive device, which acts as a gateway that we can think of as a data provider for the Octocloud ecosystem, can communicate over 4 different channels, and the data frequency of each of these channels can be changed parametrically. For example, the first channel receives data with a 5-second interval, while the second channel receives data with a 60-second interval. The main reason for this is that for some data types read, the data obtained over 60 seconds or more data extraction frequency is not meaningful. Pressure values are the best example that can be given in this regard because pressure values can change very quickly. It can rise and increase in seconds, in which case low data extraction frequency

may be needed. As a result, data extraction frequency parameters were recorded in order to obtain duration information from bool data types, and the running time of the bool data was calculated by taking the arithmetic average of the available values in the relevant summary time periods.

Consider for example the light data in an air-conditioning environment as an example. Suppose there is a light controller used for lighting the environment. The electronic component that is output while controlling the light is called a relay. The data type of the relay component that is controlled continues to exist as a bool data in the Octocloud ecosystem. It keeps the light information on the system if it is true and closed if it is false. This value is expressed as a light relay in jargon in the environment where the study is carried out. In our example, the light relay information is accumulated in our system as true and false values, but with the variable number of data we obtain, the information on how long the light remains on cannot be reached. This information was obtained by metadata extraction for bool data types in the Summary microservice. At this point, bool values are kept in two formats as summary data.

- Activation number; This counts each transition of a bool value from false to true in the system and keeps the total.
- Operation time; This value is calculated with the arithmetic to mean over the data extraction frequency information, which is variable over time, and the values of bool data.

In summary, with the Summary microservice, which is a new microservice developed, data in different data types obtained with different data retrieval frequencies are converted into a common format that can be used more easily.

### **3.1.2. Analyzing Data**

The data is processed in the summary microservice, metadata is extracted and kept in a relational database. When the resulting data are examined with reporting tools, it is understood that the necessary environment for making the estimations targeted within the scope of the thesis exists only in one facility. In cold storage facilities, each room/storage must have an energy meter to make room-based estimation and then compare it with

actual values. The controller data that controls the supply of the cooling fluid needed by the room/storage in the cooling system is known as the expansion valve relay. In split cooling systems, the data of the expansion valve is passed as the compressor relay in the system in order to provide the cold fluid from the compressor.

It is known that 60-70% of the energy consumption of a facility air-conditioned with central cooling is realized by the central cooling unit. Evans et al. (2014) In the study, the energy consumed by the central cooling unit of the facility, which is on the central cooling unit panel as the target value, is recorded and the data is kept in the system. Our input values are the metadata extracted by the summary micro-service of the expansion valve relays of the controllers that control the evaporators that physically provide the cooling of the facility. In the study, the open times of the expansion valve relays, which were extracted from the bool data, were used as input.

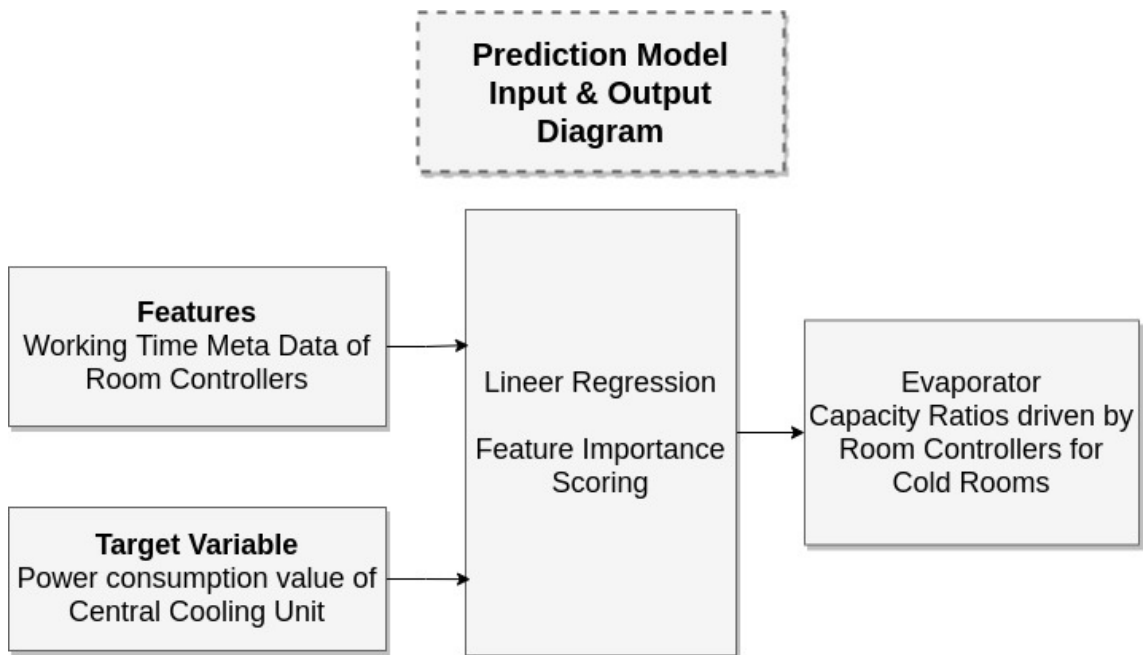
### **3.1.3. Data Wrangling**

There are data interruptions due to a problem arising from the internet and network system of the facility where the work can be done. As a result of the analysis, it was understood that at most 70% of the data that should be taken from this facility, 60 pieces per minute, 1440 pieces per day, reached the system. With the researches, it has been understood that the missing data will not seriously affect the learning stage for estimation, since the missing data is cut in a random period. In the Feature Selection study, the supervised method was applied and the Feature Importance Method, one of the filter methods, was applied with different algorithms.

The lack of data caused difficulties in manually converting the dataset to JSON format while scoring feature importance. When metadata was pulled from the database in order for the data to coincide with the same time periods, it was understood that the missing data were aligned in different time periods. This situation caused the scores to be obtained incorrectly and this situation was revealed by the analyzes made. The JSON document, which was created in order to be given to the Python modules to be scored in order to be able to score, was re-issued as a result of a more comprehensive study. As a result, it was observed that consistent scores were obtained as a result when the model was trained with the reconstructed training data set.

### 3.2. Train The Model

With the study, software consisting of several modules was developed using the Python software language. This software works with the method of giving the expansion valve relay values of the autonomous controller in each room/storage in cold storage facilities to the model as the input value. The expansion valve relay values used here are not raw data, they are the data processed by the summary microservice and kept as metadata as a result. The running times in the metadata obtained from the expansion valve bool values, Feature Importance Scoring, constitute the inputs of the training data set of the model to be made.



**Figure 3.3.** Prediction Input Output Diagram

The target value is the measurement values of the energy meter located in the central cooling unit, which is the central cooling fluid source of the facility. Since the energy consumption data mentioned here is cumulative, it cannot be used as training data as it is. For this reason, the summary microservice has metadata created in three different ways: 5 minutes, 60 minutes, and daily. In the metadata of the energy consumption values obtained, there are calculated values of the amount consumed at 5 minutes, 60 minutes, and daily intervals. The energy consumed in these intervals is included in the training dataset as the target value of the model's training dataset.

Some of the Regression methods used while training the model are as follows;

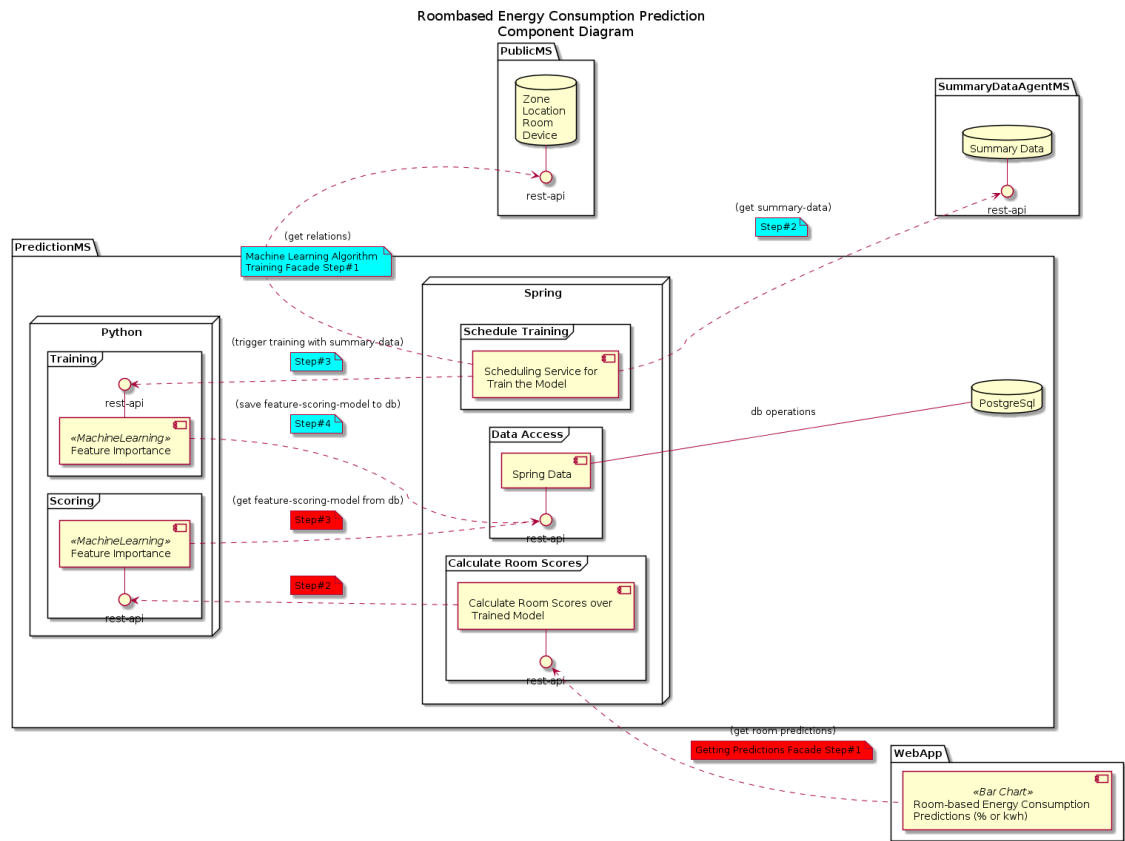
- RandomForestRegressor
- DecisionTreeRegressor
- LinearRegression

### **3.3. Test the model**

The trained model will be used to calculate the weights of the evaporators with room/warehouse or room/warehouse cooling components required to make room/warehouse based energy consumption estimation of the facilities. There are hundreds of different types and types of evaporators used in industrial cooling facilities. Each type of evaporator has a BTU value and this information is not available in monitoring systems where calculations, visualizations will be made. At this point, when calculating room/warehouse-based energy consumption, the capacity ratios of the evaporators for the facility are needed. The data set we have has been divided into parts and used as a training and test data set.

### **3.4. Deployment**

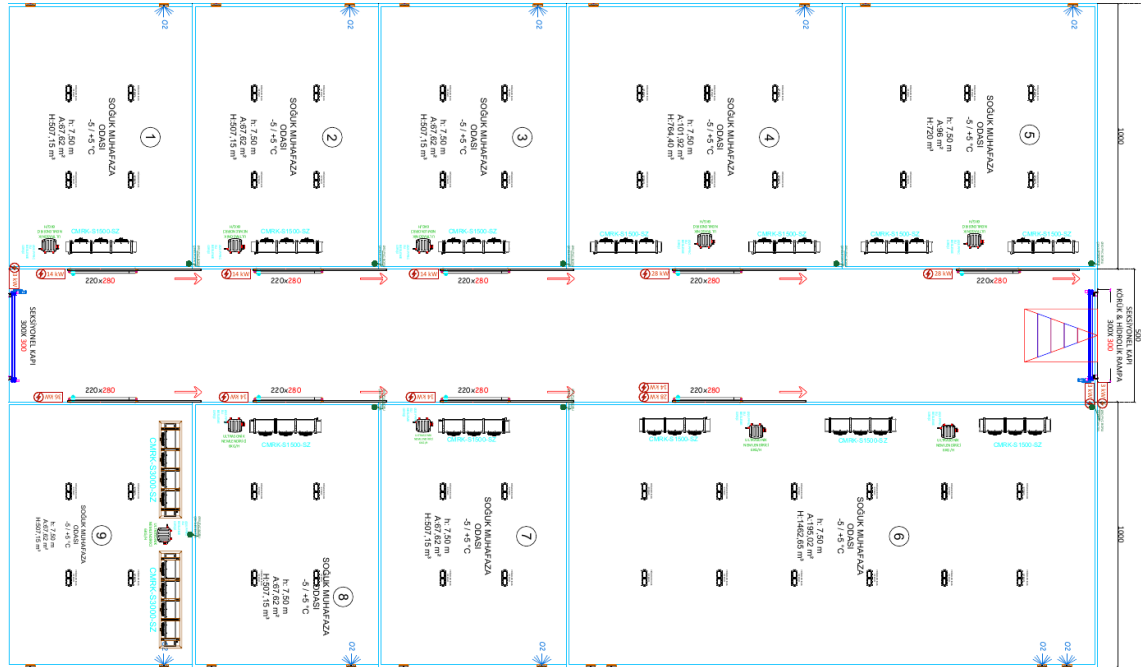
The resulting software will be deployed as a module to run inside a microservice of the Octocloud Cloud Monitoring System software as a module to calculate evaporator weights. The component diagram of the resulting software is as follows.



**Figure 3.4.** Component Diagram of Room-based Consumption Prediction

#### 4. RESULTS AND DISCUSSION

The sketch of the pilot facility, the rooms in the facility and the layout, and capacities of the evaporators in the rooms are as follows;



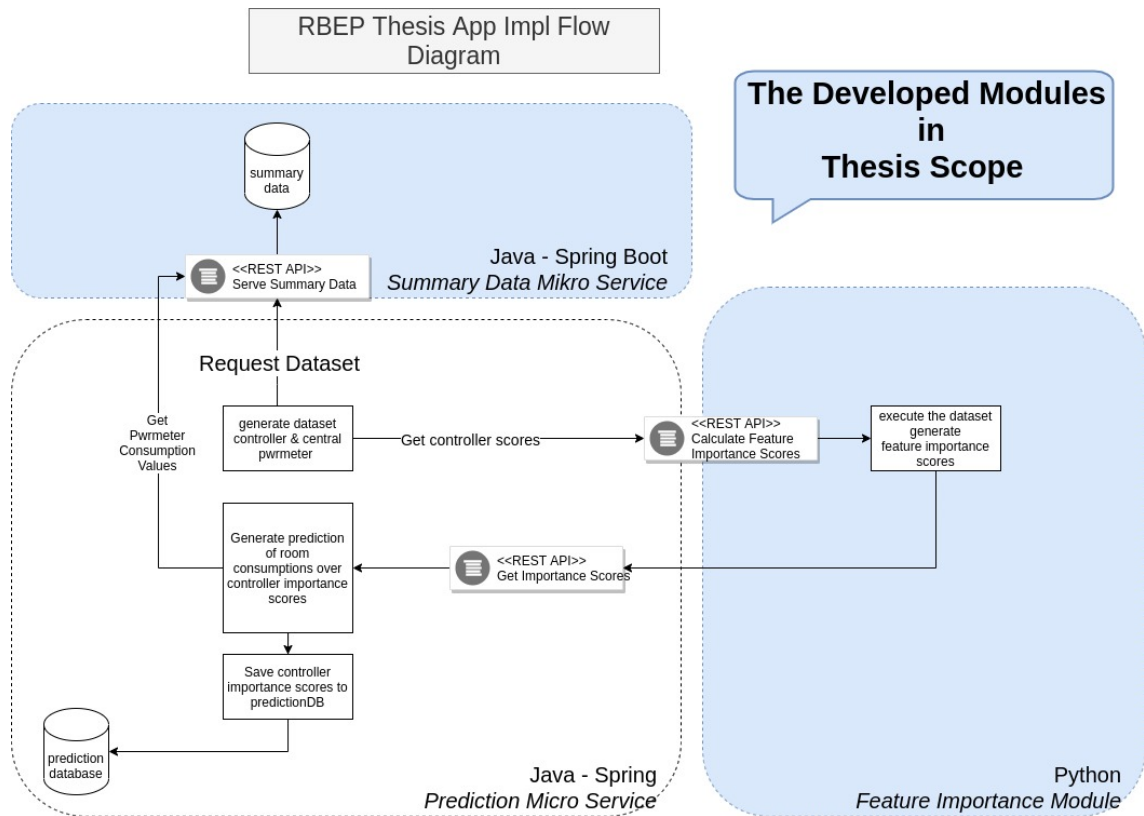
**Figure 4.5.** Pilot Facility Plan

The expansion valve times of one or more controllers located in the rooms or warehouses of the industrial cooling facility designated as the pilot zone are used as input values in the training set. The data measured by the power meter device, which is located on the panel of the central cooling unit and measures the consumption of the central cooling system, is used as the target value of the training data set. In the studies, a trained model was obtained based on the facility. It is aimed to obtain the weights of the evaporators over the trained model using the data set with new data.

To make room consumption predictions on the basis of the targeted room or warehouse, the capacities of the evaporators must be known or predicted. The subject of the thesis is to calculate the ratios of the capacities of one or more evaporator components in the rooms or warehouses of the cooling facilities with each other by using Regression methods to predict room consumption. To fulfill this hysteria, the following Regression methods are used for Feature Selection.

Some of the Python scikit-learn library's regression methods used while training





**Figure 4.6.** Microservices Flow Diagram

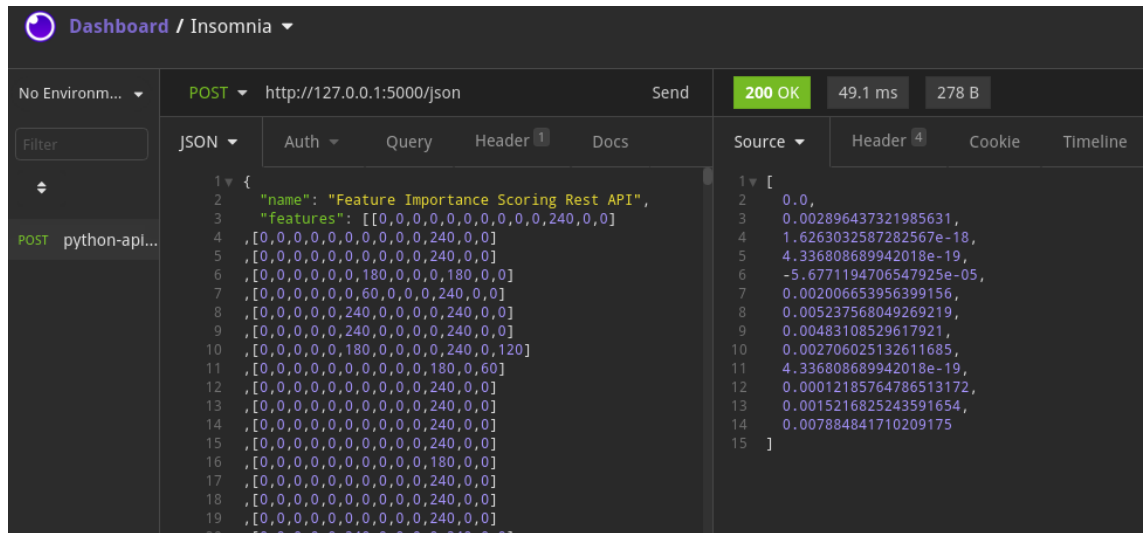
the model are as follows;

- Linear Regression
- Random Forest Regressor
- Decision Tree Regressor

#### 4.1. Results

The example is based on 8600 5-minute summary data of 30 days. The study was carried out with a data set with 70% of the data expected to be due to technical problems. Sending the mentioned data set to RESTful API to extract the feature score and the sample response is shown below.

In the table below, the controllers that manage the rooms in the pilot facility and which evaporators these controllers manage are given in detail. The aim of this thesis covers the systems with central cooling, as explained in detail in the introduction and summary sections, so the data specified as cooling type split in the table are ignored.



**Figure 4.7.** Feature Importance Scores - Result of RESTful API

	Cooling Type	Controller Type	Working Time	Evaporator 1	Evaporator 2	ROOM Total BTU Capacity (Central System)
Freeze Room	Central	FREEZE	0 min	1500BTU		1500 BTU
Room 1	Central	VPM	21 hr	1500BTU	-	1500 BTU
Room 2	Central	VPM	0 min	1500BTU		1500 BTU
Room 3	Central	VPM	0 min	1500BTU		1500 BTU
	Split	JR	54 hr	1500BTU		
Room 4	Central	VPM	139 hr	1500BTU		1500 BTU
Room 5	Central	VPM	27 hr	1500BTU	1500BTU	3000 BTU
Room 6	Central	VPM	66 hr	1500BTU	1500BTU	4500 BTU
	Central	VPM	70 hr	1500BTU		
Room 7	Central	VPM	0 min	1500BTU		1500 BTU
	Split	JR	401 hr	1500BTU		
Room 8	Central	VPM	165 hr	1500BTU		1500 BTU
Room 9	Central	VPM	82 hr	3000 BTU	3000 BTU	6000 BTU

**Figure 4.8.** Facility Technical Details

The types of controllers used in the facility can manage a maximum of two evaporators. Each evaporator has a cooling capacity and this capacity is specified as BTU value. The BTU numbers are just represents the ratio of the evaporator capacities driven by controllers for rooms. Room controller evaporator and evaporator capacity information are included in the table. specified in the time column.

Metadata obtained from the data obtained from the controllers in the pilot facility are kept in the Summary microservice. The summary has been developed within the scope of microservice thesis studies. The data obtained from the relevant API of this service is given as a training data set to the relevant API in the module where the feature selection

scoring API is located, and the results are received as a response. The results obtained with three different regression methods are given in the table below.

	Cooling Type	Controller Type	Working Time	Evaporator 1	Evaporator 2	Lineer Regression	Randam Forest	DecisionTree
Freeze Room	Central	FREEZE	0 min	1500BTU		0.0000	0.0000	0.0000
Room 1	Central	VPM	21 hr	1500BTU	-	0.0029	0.0261	0.0266
Room 2	Central	VPM	0 min	1500BTU		0.0000	0.0000	0.0000
Room 3	Central	VPM	0 min	1500BTU		0.0000	0.0000	0.0000
	Split	JR	54 hr	1500BTU		-0.0001	0.0148	0.0160
Room 4	Central	VPM	139 hr	1500BTU		0.0020	0.0500	0.0511
Room 5	Central	VPM	27 hr	1500BTU	1500BTU	0.0052	0.0615	0.0599
Room 6	Central	VPM	66 hr	1500BTU	1500BTU	0.0048	0.2610	0.2666
	Central	VPM	70 hr	1500BTU		0.0027	0.0238	0.0205
Room 7	Central	VPM	0 min	1500BTU		0.0000	0.0000	0.0000
	Split	JR	401 hr	1500BTU		0.0001	0.0308	0.0254
Room 8	Central	VPM	165 hr	1500BTU		0.0015	0.0311	0.0285
Room 9	Central	VPM	82 hr	3000 BTU	3000 BTU	0.0079	0.5009	0.5054

**Figure 4.9.** Regression - Feature Importance Scores - Evaporation Weights

It has been observed that the Linear Regression method is the most successful method among the methods used in feature importance scoring above.

As we understand from the relevant design plan of the pilot facility, Room-9 has a 3000BTU + 3000BTU total capacity of 6000 BTU. Room-8 has a 1500BTU capacity. It is understood that there is a 1/4 ratio between the cooling capacity of the two rooms.

When the results obtained are examined, it is understood that the closest estimation to the real capacities is made by the Linear Regression method.

**Table 4.6.** Ratio Prediction Results for Linear Regression

<b>Controller</b>	<b>Actual Capacity Ratio</b>	<b>Predicted Ratio</b>	<b>Error</b>	<b>Squared Error</b>
Room 4 Controller 1	10	10	0	0
Room 5 Controller 1	20	26	6	36
			MAE: 3	MSE: 18
Room 4 Controller 1	10	10	0	0
Room 6 Controller 1	20	24	4	16
			MAE: 2	MSE: 8
Room 8 Controller 1	10	10	0	0
Room 9 Controller 1	40	52	12	144
			MAE: 6	MSE: 72

## 5. CONCLUSION

In this study, the prediction of the consumption of the rooms in the central cooling system, which is the common method used in the industrial cold storage area, was carried out. The energy consumption of the rooms or warehouses in the facilities is not known. In order to calculate this consumption, it is aimed to use the power meter data on the board of the central cooling unit, which is the component of the central cooling system that supplies the cold flows.

The rooms or warehouses in the facilities are controlled by one or more autonomous controllers. The factor that determines the number of controllers is the BTU capacity needed to cool the rooms. If the ideal BTU capacity can be provided with how many evaporators, as many evaporators are deployed in the room or warehouse. The biggest challenge faced within the scope of the study was to predict the total BTU capacity of the rooms and autonomous controllers. To eliminate this uncertainty, Machine Learning Linear Regression methods were used and the weights of the evaporators in the facility to each other were tried to be predicted over the model trained with the training data set.

As a result of the thesis work, metadata generation from raw data and evaporator BTU capacity prediction modules were developed. By using these modules, the consumption of the rooms can be calculated over the data of the central cooling unit and autonomous room controller.

The thesis is also an RandD Center Project. By using these outputs in the relevant project, consumption data on the basis of room or warehouse can be presented to users as a feature in facilities cooled by the central system.

## 6. REFERENCES

- Ahmad, A., Hassan, M., Abdullah, M., Rahman, H., Hussin, F., Abdullah, H., and Saidur, R. 2014. A review on applications of ann and svm for building electrical energy consumption forecasting. *Renewable and Sustainable Energy Reviews*, 33, 102–109. <https://doi.org/https://doi.org/10.1016/j.rser.2014.01.069>
- Amasyali, K., and El-Gohary, N. M. 2018. A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, 81, 1192–1205. <https://doi.org/https://doi.org/10.1016/j.rser.2017.04.095>
- Ayan, S. M. İ. 2018. *Power estimation of photovoltaic system with the artificial intelligence algorithm*, Master's thesis. Kırklareli Üniversitesi. Kırklareli.
- Azadeh, A., Ghaderi, S., and Sohrabkhani, S. 2008. Annual electricity consumption forecasting by neural network in high energy consuming industrial sectors. *Energy Conversion and Management*, 49, 8., 2272–2278. <https://doi.org/https://doi.org/10.1016/j.enconman.2008.01.035>
- Dandıl, E., and Gürgen, E. 2019. Yapay sinir ağları kullanılarak fotovoltaiik panel güç çıkışlarının tahmini ve sezgisel algoritmalar ile karşılaştırılması. *Avrupa Bilim ve Teknoloji Dergisi*, 146–158. <https://doi.org/10.31590/ejosat.540262>
- Pirim, H. 2006. Yapay zeka. *Yaşar Üniversitesi E-Dergisi*, 1, 81–93. <https://doi.org/10.19168/jyu.72783>
- Dong, Z., Liu, J., Liu, B., Li, K., and Li, X. 2021. Hourly energy consumption prediction of an office building based on ensemble learning and energy consumption pattern classification. *Energy and Buildings*, 241, 110929. <https://doi.org/10.1016/j.enbuild.2021.110929>
- Einstein, A. 1939. On a stationary system with spherical symmetry consisting of many gravitating masses. *Annals of Mathematics*, 40, 4., 922–936. <http://www.jstor.org/stable/1968902>
- Elmas, Ç. 2007. *Yapay zeka uygulamaları: Yapay sinir ağları - makine öğrenmesi - derin öğrenme - derin ağlar - bulanık mantık - sinirsel bulanık mantık - genetik algoritma*. Seçkin Yayıncılık.
- Erler, M., Sağıroğlu, Ş., and Beşdok, E. 2003. *Mühendislikte yapay zeka uygulamaları 1 / yapay sinir ağları*. Ufuk Kitapevi.

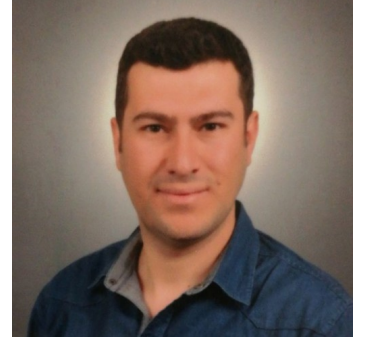
- Evans, J. [J.A.], Foster, A., Huet, J.-M., Reinholdt, L., Fikiin, K., Zilio, C., Houska, M., Landfeld, A., Bond, C., Scheurs, M., and van Sambeek, T. 2014. Specific energy consumption values for various refrigerated food cold stores. *Energy and Buildings*, 74, 141–151. <https://doi.org/10.1016/j.enbuild.2013.11.075>
- Evans, J. [JA], and Gigiel, A. 2007. Reducing the energy consumption in cold stores. *The 22nd International Congress of Refrigeration, 21-26 August 2007*.
- Evans, J. [J.A.], Hammond, E., Gigiel, A., Fostera, A., Reinholdt, L., Fikiin, K., and Zilio, C. 2014. Assessment of methods to reduce the energy consumption of food cold stores. *Applied Thermal Engineering*, 62, 2., 697–705. <https://doi.org/10.1016/j.applthermaleng.2013.10.023>
- Haenlein, M., and Kaplan, A. 2019. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61, 4., 5–14. <https://doi.org/10.1177/0008125619864925>
- Kingdom, E. U. *Energy consumption guide 37 : Cold storage sector*. Energy Efficiency Office, Department of the Environment. 1994, February.
- Mcculloch, W., and Pitts, W. 1943. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 127–147.
- Orosz, J. A., McClintock, J. E., Aufdenberg, J. P., Remillard, R. A., Reid, M. J., Narayan, R., and Gou, L. 2011. The mass of the black hole in cygnus x-1. 742, 2., 84. <https://doi.org/10.1088/0004-637x/742/2/84>
- Özlu, G. 2018. *Electricity consumption data analysis of kayseri and vicinity using artificial bee colony algorithm*, Master's thesis. Erciyes Üniversitesi. Kayseri.
- Öztemel, E. 2012. *Yapay sinir ağları*. Papatya Bilim Yayıncılık.
- Peter Timmer, C. 1988. Chapter 8 the agricultural transformation. Elsevier. [https://doi.org/10.1016/S1573-4471\(88\)01011-3](https://doi.org/10.1016/S1573-4471(88)01011-3)
- Shaw, K. 2008. The application of artificial intelligence principles to teaching and training. *British Journal of Educational Technology*, 39, 319–323. <https://doi.org/10.1111/j.1467-8535.2008.00817.x>
- Utku, T. 2015. Storage is the key. *Cantek Foundation for Science and Culture*, 14., 1–72.
- Werner, S., Vaino, F., Merts, I., and Clelenad, D. 2006. Energy use in the new zealand cold storage industry. *IIR-IRHACE Conference, The University of Auckland*.

- Wu, D., Shen, J., Tian, S., Zhou, C., Yang, J., and Hu, K. 2018. Experimental study of temperature characteristic and energy consumption of a large-scale cold storage with buried pipe cooling. *Applied Thermal Engineering*, 140, 01–53. <https://doi.org/10.1016/j.applthermaleng.2018.05.035>
- Zhao, H.-x., and Magoulès, F. 2012. A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 16, 6., 3586–3592. <https://doi.org/https://doi.org/10.1016/j.rser.2012.02.049>



## CURRICULUM VITAE

Burak ERKAN  
burak@yazilim.vip



### EDUCATION DETAILS

Graduate 2018-2021	Akdeniz University Institute of Natural Sciences, Department of Computer Engineering, Antalya
Undergraduate 2004-2010	Cyprus International University Engineering Faculty, Department of Computer Engineering, Nicosia

### WORK EXPERIENCES

Yazılım VIP 2013-present	Co-founder
Cantek Group 2015-2021	Software Development Supervisor