VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY

UNIVERSITY OF INFORMATION TECHNOLOGY

FACULTY OF COMPUTER ENGINEERING

NGUYEN MANH THAO

QUACH THE HAO

CAPSTONE PROEJECT

LOCAL WEATHERMONITORING, FORECASTING SYSTEM USING NEURAL NETWORK AND LORA COMMUNICATION TECHNOLOGY

THIẾT BỊ QUAN TRẮC, DỰ BÁO THỜI TIẾT PHẠM VI NHỎ SỬ DỤNG CÔNG NGHỆ MÁY HỌC VÀ GIAO TIẾP TRUYỀN THÔNG LORA

BACHELOR OF COMPUTER ENGINEERING

HO CHI MINH CITY, 2019

VIETNAM NATIONAL UNIVERSITY HO CHI MINH CITY

UNIVERSITY OF INFORMATION TECHNOLOGY

FACULTY OF COMPUTER ENGINEERING

NGUYEN MANH THAO – 14520853

QUACH THE HAO – 14521124

CAPSTONE PROEJECT

LOCAL WEATHERMONITORING, FORECASTING SYSTEM USING NEURAL NETWORK AND LORA COMMUNICATION TECHNOLOGY

THIẾT BỊ QUAN TRẮC, DỰ BÁO THỜI TIẾT PHẠM VI NHỎ SỬ DỤNG CÔNG NGHỆ MÁY HỌC VÀ GIAO TIẾP TRUYỀN THÔNG LORA

BACHELOR OF COMPUTER ENGINEERING

MENTOR

PhD. TRINH LE HUY

HO CHI MINH CITY, 2019

LIST OF DEFENSE COMMITTEE

Defense committee established under Decision No ………………………………. dated on……………………. of Rector of the University of Information Technology.

1. – .... Chairman.
2. – .... Secretary.
3. – .... Commissioner.

(Nhận xét của giảng viên hướng dẫn)

(Nhận xét của giảng viên phản biện)

ACKNOWLEDGEMENTS

Foremost, we would like to express our sincere thanks to teachers of University of Information Technology in general and the Faculty of Computer Engineering for imparting knowledge and valuable experience during the past 5 years to us.

We wish to send our sincere thanks to Mr. Nguyen Minh Son, Dean of the Faculty of Computer Engineering, for providing us with all the necessary facilities for the research.

We are grateful to Mr. Trinh Le Huy, lecturer, in the Faculty of Computer Engineering. We are extremely thankful and indebted to him for sharing expertise, and sincere and valuable guidance and encouragement extended to us.

We deeply thanks to the help of our friends, seniors who helped us to find information during the course.

We are also grateful to our parents for the unceasing encouragement, support, and attention.

Once again, we would like to thank all those who are interested in our thesis. However, in the process of working there may be some unavoidable mistakes, we are looking forward to receiving comments on this thesis.

Nguyen Manh Thao - Quach The Hao

VIETNAM NATIONAL UNIVERSITY **UNIVERSITY OF INFORMATION TECHNOLOGY**	**SOCIALIST REPUBLIC OF VIETNAM** **Independence - Freedom - Happiness**

OUTLINE DETAILS

PROECT NAME: LOCAL WEATHERMONITORING, FORECASTING SYSTEM USING NEURAL NETWORK AND LORA COMMUNICATION TECHNOLOGY
Mentor: Ph.D Trinh Le Huy
Duration: 03/09/2018 - 01/01/2019
List of students: **Nguyen Manh Thao– 14520853** **Quach The Hao – 14521124**
Contents: **Thesis target** Nowadays, along with the development of the Internet of Things models in industries and services, the demand for weather forecasts is increasing. However, there is currently no simple system to meet this need. The current solution to this problem is using National weather forecasting channels. An advantage of this solution is that it is no need for additional hardware. It is entirely free because forecast information is published on media channels. However, there is a disadvantage that the forecast information from National weather forecasting

channels is not completely correct for all local areas (where the public measurement devices aren’t installed) .

The thesis "Local weather forecasting system using neural network" focus on forecasting the weather for a small localized area using a neural network with a reasonable cost, personalizing for users, serving for family and small businesses.

Thesis focuses on researching and building a system that can collect environmental data and forecast the precipitation in a small-scale area using Machine Learning. Utilizing the advantages of the Machine Learning algorithm is to improve predictive accuracy along with the working time. Applying communication technology - LoRa in transmit and receive operations to increase the device installation distance and save power for monitoring nodes.

Thesis scope

Build an environmental monitoring and weather forecasting system with high accuracy in local areas. The system must be easy to install, the accuracy of forecasting results increases along with system use time. Detail:

- Collect environmental data in 3 months at least. Process data and use it to train the Back-propagation Neural Network model in the beginning.
- Develop a local weather forecasting application using Back-propagation Neural Network model. The application has 7 inputs for forecasting include temperature, humidity, atmosphere, the hour in the day, month in the year, wind direction and wind speed. For training operation, the application will use rainfall.
- Communicate and get data from sensors
- Build a power supply system with Li-ion batteries and use a solar panel to maintain battery life
- Communicate environmental monitoring devices and a gateway with LoRa technology

- Develop a website with HTML, CSS, and NodeJS. The website must have an easy-to-use interface and stable operation.
- Develop a database system managing sensor data, forecasting results and user's data with MongoDB

Thesis objects

- Weather Station with Anemometer/Wind vane/Rain bucket SKU:SEN0186
- Solar panel, Li-ion battery and battery charger circuit.
- Mesh-protected Weather-proof Temperature/Humidity Sensor
- Barometric Pressure/Temperature/Altitude Sensor BMP180
- Theoretical basis in weather forecasting field
- Machine Learning Back-propagation Neural Network model
- LoRa Transceiver RFM96
- Wi-Fi Module ESP8266
- MQTT Protocol
- Web Server with NodeJS programming language
- MongoDB Database

Research methods

- Research forecasting techniques and factors that affect the amount of rainfall, especially for Vietnam
- Develop a backpropagation neural network model for forecasting the amount of rainfall in a small-scale area
- Research and develop a system that collect environmental values as inputs of the neural network model using LoRa communication technology

Expected results

- Functioning system and packaged nodes to be able to work outdoor
- Increase rain forecasting accuracy to 63% comparing to the system in 2016
- Design the anti-collision algorithm to decrease packet collision rate

Plan:

Week No.	Contents	Report date
1 (3/9 - 8/9)	Research weather forecasting related papers and theories	09/09/2018
2 (10/9 - 15/9)	Check and develop libraries for wind, temperature, humidity and atmosphere sensors	16/09/2018
3 (17/9 - 22/9)	Design the system overview	23/09/2018
4 (24/9 - 29/9)	Make monitoring node and gateway hardware prototypes	30/09/2018
5 (1/10 - 6/10)	Packaging monitoring node for outdoor operating	07/10/2018
6 (8/10 - 13/10)	Develop monitoring node and gateway firmware	14/10/2018
7 (15/10 - 20/10)	Collect data for training ANN model	21/10/2018
8 (22/10 - 27/10)		28/10/2018
9	Database handler	04/11/2018

(29/10 - 3/11)		
10 (05/11 - 10/11)	MQTT Broker	11/11/2018
11 (12/11 - 17/11)	Solving collision problem on LoRa bandwidth	18/11/2018
12 (19/11 - 24/11)	Optimize power supply module for node	25/11/2018
13 (26/11 - 1/12)	Develop website showing collected data and forecast result	02/12/2018
14 (3/12 - 8/12)	Checking and fixing bugs	09/12/2018
15 (10/12 - 15/12)	Checking and fixing bugs	16/12/2018

Detail work:

Student	Work
Nguyen Manh Thao	– Build basic system, collect sample data for 3 months to build and model machine learning training – Communicate and read data from sensors. – Build a power supply for data collection systems with li-ion batteries, use solar batteries to maintain battery operation – Perform communication between data receiving station and central board with LoRa wave.

	– Build a system management Web site – Packaging products – Check and fix errors
Quach The Hao	– Research on theoretical basis for weather forecast and data – Develop Neural Network models to predict the weather using 7 input variables including temperature, humidity, pressure, hour of day, month of year, wind direction and wind force. – Gateway box design – Collect predictive results, give a performance comparison table between models

Confirm of Mentor (Signature and Full name) **PhD. Trinh Le Huy**	**HCMC, Jan 9th, 2018** **Student** (Signature and Full name) **Nguyen Manh Thao**

MENU

FIGURE MENU

TABLE MENU

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
VPS	Virtual Private Server
ARIMA	Autoregressive Integrated Moving Average
SPI	Serial Peripheral Interface
GPIO	General Purpose Input/output
RDBMS	Relational Database Management System
MAC	Medium Access Control
RTC	Real Time Clock
OTA	Over the Air

ABSTRACT

According to the statistics of 2017, the structure of Viet Nam's economy is as follows: agriculture, forestry, and aquaculture account for 15.34%; industrial and construction area account for 33.34%; service area accounts for 41.32%. In particular, agriculture, forestry, and aquaculture, construction play essential roles, occupying a significant share and now they are still growing strongly. Most of the models in these industry areas are affected by the weather, and local weather forecasts are of great benefit. Along with the development of the Internet of Things models in the industry, the demand for weather forecasts is become very important. However, there is currently no simple system to meet this need.

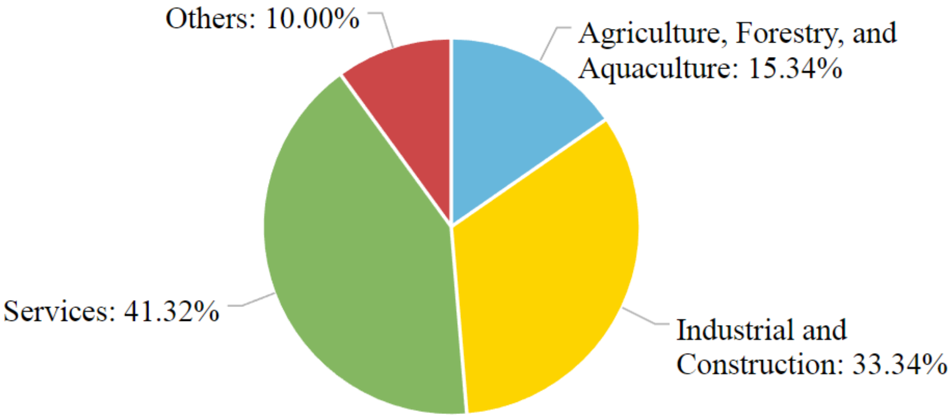


Figure 1: Vietnam economy structure in 2017

The current solution to this problem is using National weather forecasting channels. An advantage of this solution is that it is no need for additional hardware. It is entirely free because forecast information is published on media channels. However, there is a disadvantage that the forecast information from National weather forecasting channels is not completely correct for all local areas.

The thesis "Local weather forecasting system using neural network" focus on forecasting the weather for a small localized area using a neural network, personalizing for users, serving for family and small businesses.

The system includes a central device and other environmental monitoring devices that collect data and transmit to the central device using LoRa technology - the new wireless technology that provides the long-range transmission with a very low power consumption. The thesis uses Neural Network as the main algorithm for forecasting weather with input data collected from the environment such as temperature, humidity, atmospheric pressure, rainfall, wind direction and speed, the hour in the day, month in the year. The system also has a web server so that users can connect to when they have internet.

This thesis report contains five chapters:

Chapter 1: Overview

Chapter 2: Theory and Experimental Study

Chapter 3: System Analysis and Design

Chapter 4: Result and Summary

Chapter 5: Future work

Chapter 1. **OVERVIEW**

1.1 Background

Today, global climate change is day-by-day increasingly affecting people's lives. Climate change has caused serious disasters such as tsunamis, earthquakes, floods, etc. So, it is imperative to find solutions to prevent and minimize damage caused by those disasters.

Figure 2: Flash flood’s result

According to statistics in 2017, the South China Sea has 16 storms and 4 tropical depression with 6 storms directly affect Vietnam. Natural disasters in 2017 left 386 dead and missing, 654 injured; More than 8000 houses collapsed; Total economic losses is about 60000 billion VND. Particularly, the typhoon No. 12, which landed on November 4, 2017, caused losses of more than 1300 billion VND in Khanh Hoa province.

Figure 3: Weather Monitoring Station at Tan Son Nhat airport

The reason is that Vietnam is located in the tropical monsoon climate region, Pacific belt, next to the South China Sea is, it's inevitable that the storm often occurs. In addition, rivers in Vietnam is often short and steep, when it's heavy rain, the water rises rapidly, causing landslide and flash floods. Besides, there are many activities in Vietnam that depend on weather forecasting information like agriculture, fisheries, construction, etc. For example: In construction, repairing the building after a sudden rain will increase the cost of construction as well as reduce the quality. Therefore, monitoring the weather before construction is very important and there are many companies pay attention to. With a project under construction in Thu Duc district, based on weather forecasting results on the Web site of the National Hydro - Meteorological Forecasting Center (http://www.nchmf.gov.vn), it will not rain at the time of pouring concrete. However, the weather information is only collected at Weather Monitoring Station located at Tan Son Nhat Airport (10°49'01.2"N 106°40'01.2"E), the results are only accurate for the area around Tan Binh district. At nearly 20 km, the accuracy of this result is no longer guaranteed. Therefore, the need for a local weather forecasting system is extremely urgent.

1.2 Related work

1.2.1 International research

Predicting the weather has long been a common problem for mankind. There have been many advances in meteorological and hydrographic science and invent modern equipment. But local weather forecasting is not popular. Most local weather forecasting devices in the market used to notify the weather state or get weather information from the Internet. And they also have the same disadvantages as solutions mentioned above.

The article [1] presents the method using the Artificial Neural Networks model with the supervised learning paradigm to predict the mean daily temperature values with the input data for the meteorological station located in Bangkok. The Artificial Neural Network model was trained by the Fireworks algorithm - the ANN optimization algorithm developed by the Swarm Intelligence Algorithm. However, the article focuses on algorithm development and research rather than product development.

In 2013, Andrew Culclasure has published the Master thesis [2] about using Neural Networks in Local Weather Forecasts field at Georgia Southern University. The thesis focuses on performing a survey on Weather forecasting models using ANN. In addition, the subject also introduced an author's experiment on the use of ANN for predicting temperature. However, the topic is focused on surveying.

In 2012, Author Kumar Abhishek published an article [3] on Procedia Technology magazine Volume. In Pages 311-318, the article presents considering the possibility of the method that uses ANN with different transmission parameters, different number of hidden layers and neurons in the weather prediction, especially the maximum temperature in a year.

1.2.2 Domestic research

Vietnam is located in an area where the weather is complex, frequent floods occur, especially Central and South Vietnam. However, most meteorological stations in Vietnam are large scale and not developed in each locality. Pieces of equipment are imported from abroad with high cost and the accuracy is not high due to the large scope of monitoring. Besides, the number of the meteorological station is quite a few.

In Vietnam Maker Contest with Intel Galileo 2015 (VMIG 2015), the team from Hanoi University of Science won the 3rd prize with a project named "Trạm quan trắc môi trường và cảnh báo sớm thiên tai" (Environmental monitoring station and warning of natural disasters). The system includes 3 main components: sensors, central processing board which is Intel Galileo and a website. This research is highly appreciated by the judges, but the system still has some disadvantages that it cannot predict natural disaster in advance.

Mimosatek company won the 1st prize in Venture Cup Contest 2015 with a product named "Hệ thống quản lý tưới chính xác", a device used to measure environmental values, process the information and calculate the amount of water the plant needs. All of the processes are managed and controlled via smartphones. But this system just measures environmental values at that point to water the plant reasonable and warn the users. It cannot predict the weather in advance.

1.3 Target, object and scope of the thesis

1.3.1 Target and object of this thesis

Our thesis focuses on researching and upgrading the previous thesis [4] in 2016. Build a system that can collect environmental data and forecast without using third-party data. Environmental monitoring devices can operate for long periods of time outside of the actual environment. Package the product so that the device can normally operate under the rain and wind.

The thesis uses Machine Learning, specifically the Back-propagation Neural Network model to solve the problem of small-scale rain forecasting. Utilizing the advantages of the Back-propagation Neural Network model is to improve predictive accuracy over time.

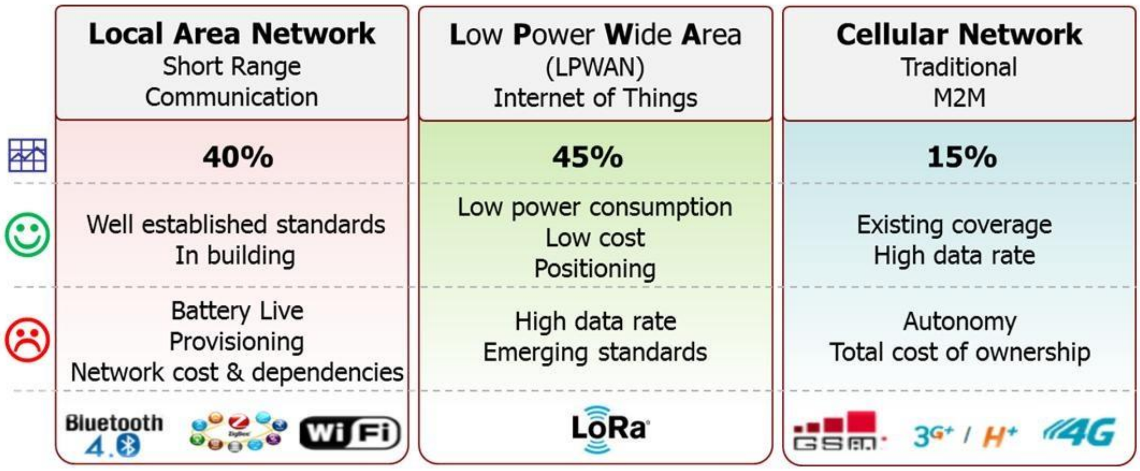


Figure 4: Comparison of LPWAN

We choose to use LoRa in this thesis because it has a long distance transmission. Therefore, we can install node on the roof or any convenient location, because the node needs to be installed at an open location to collect wind information and the gateway need to be installed where there is an internet connection. Besides, Lora technology can save more power than other transmit technologies as in Figure 4. Also, thanks to solar cells, the battery life of the monitoring devices is extended, so that the system can be installed in many different areas without other power supplies.

Using Wi-Fi module in Gateway can send weather data to Linux VPS in other to predict with Machine Learning. Also, Linux VPS run a Web Server application, provides users with an easy-to-use managing website.

Overall, the biggest problem we focus on solving is that accuracy and reliability of the system must be optimized.

Detailed target:

- Build a hardware system that collects weather data independently. The hardware must be packaged to able to work in windy weather environment. The hardware system includes:
    - Environmental monitoring devices: collect weather data such as temperature, humidity, atmosphere, wind direction, wind speed, and rainfall. Using a solar panel to maintain the battery life.
    - Gateway: Receive data from Environmental monitoring devices via Lora and push them to Linux VPS via Wi-Fi. Receive weather forecasting results from the server on VPS and display them on the LCD.
- Research weather forecasting theory and study papers about weather forecasting. Analyze data collected from the environment in a long time to obtain a theoretical basis for weather forecasting.
- Research and develop a local weather forecasting application using Back-propagation Neural Network model. Compare it to the local weather forecasting application using Fuzzy Logic. Increase forecasting accuracy to 63%.

- Develop a system managing sensor data, forecasting results and user's data with MongoDB
- Develop a Web Server on Linux VPS with an easy-to-use interface and stable operation.

System design:

- Weather Station with Anemometer/Wind vane/Rain bucket SKU:SEN0186
- Solar panel, Li-ion battery and battery charger circuit.
- Mesh-protected Weather-proof Temperature/Humidity Sensor
- Barometric Pressure/Temperature/Altitude Sensor BMP180
- Theoretical basis in weather forecasting field
- Machine Learning Back-propagation Neural Network model
- LoRa Transceiver RFM96
- Wi-Fi Module ESP8266
- MQTT Protocol
- Web Server with NodeJS programming language
- MongoDB Database

1.3.2 Scope

Build an environmental monitoring and weather forecasting system with high accuracy in local areas. The system must be easy to install, the accuracy of forecasting results increases along with system use time.

Detail:

- Collect environmental data in 3 months at least. Process data and use it to train the Back-propagation Neural Network model in the beginning.
- Develop a local weather forecasting application using Back-propagation Neural Network model. The application has 7 inputs for forecasting include temperature, humidity, atmosphere, the hour in the

day, month in the year, wind direction and wind speed. For training operation, the application will use rainfall.

- Communicate and get data from sensors
- Build a power supply system with Li-ion batteries and use a solar panel to maintain battery life
- Communicate environmental monitoring devices and a gateway with LoRa technology
- Develop a website with HTML, CSS, and NodeJS. The website must have an easy-to-use interface and stable operation.
- Develop a database system managing sensor data, forecasting results and user's data with MongoDB

1.4 Advantage and disadvantage

1.4.1 Advantage

During the period of the thesis, we have received the enthusiastic support of our relatives, family, friends and especially thesis mentor. Besides, we also have received a lot of support from the Faculty of Computer Engineering, such as the equipment, research tools, working rooms and teachers in the department helped us in finding a lot of useful materials as well as solutions to complete the thesis. The University of Information Technology Facilities Management Department also helped a lot by permitting us to put environmental monitoring devices at the university campus. The collaboration between members of our team makes a significant contribution to the thesis success. We have a good plan for this thesis and a sense of responsibility. Also, the sharing of knowledge and experience of other research teams is also a resource for our team to complete this thesis.

1.4.2 Disadvantage

Besides the advantages mentioned above, the group has also encountered difficulties during the period of the thesis. Typically, the lack of expertise in meteorology or knowledge we found is not exactly accurate. This affects the accuracy of forecasting results. Moreover, we need to collect environmental data and build environmental monitoring devices at the same time. For this reason, environmental data collecting was interrupted at some specific time, which lead to a lack of environmental data. Some sensors we use in this thesis are costly, rare and hard to find in the domestic market. Building forecasting application along with Web Servers on University's Linux VPS also faces many difficulties due to resource constraints and the need for network access permissions.

Chapter 2. **THEORY AND EXPERIMENTAL STUDY**

2.1 Weather forecasting [5]

2.1.1 Theory of weather forecasting

Weather forecasting is the application of science and technology to predict the conditions of the atmosphere for a given location and time. Weather forecasting means that any advance information about the probable weather in future is obtained by evaluating the present and past meteorological conditions of the atmosphere.

In weather forecasting, the advance information of weather elements such as the distribution of rainfall, temperature changes, weather thunderstorm hailstorm, snow or frost, sky cover, winds, humidity, cloudiness, etc. is provided.

Weather forecasting based on their validity periods or timescale is classified as follows:

Short range weather forecast

Forecasting of coming weather for 2 to 3 days in advance is called short-range weather forecasting. Weather phenomena for which forecast are issued under this type is:

- Rise and fall in temperature
- Speed and direction of the wind
- Cloudiness
- Rainfall amount

This type of weather forecast is issued by India Meteorological Department, Delhi for different regions throughout the country. The error in forecast ranges from 20-30 percent.

Medium range weather forecast

Forecasting of weather 4-10 days in advance is termed as medium-range forecasting. Presently, medium range weather forecasting for 4 days in advance is issued by National Centre for Medium Range Weather Forecasting (NCMRWF) situated in NOIDA, U.P. This forecast is issued twice a week i.e. on Tuesday and Friday forecast error ranges from 30 to 40 percent. Weather phenomena on which weather forecast is issued are as follow:

- Increase or decrease in maximum and minimum temperature
- Cloudiness
- Speed and direction of wind
- Precipitation amount

Long range weather forecast

Forecasting of coming weather for more than 10 days or a month or a season in advance is called long-range weather forecasting. It may also be categorized as a monthly forecast or seasonal forecast depending on the time. Long-range forecasting is issued by India Meteorological Department, New Delhi for southwest monsoon rains, first in April and then the forecast is updated in July. This forecast is issued region is divided into four zones: (1) Northeastern region (2) Central region (3) North Western Region (4) Peninsular region. The forecasting error varied around 40 percent. The weather parameters on which forecast have been issued as under:

- The onset of monsoon rains
- Deviation of rainfall amount from long period average (LPA) rainfall

In the Agricultural Meteorological Curriculum Document (Đoàn Văn Điếm - editor) of Agricultural University 1, in chapter XI, section 1.3b [4], weather phenomena showing signs of rain and thunderstorms is mentioned as followings:

Weather phenomena show good weather [6]

- The air pressure increases constantly, or less variation in the day
- It less cloudy at night, many stars
- In the morning, there are many cumulus clouds, flat rattan legs, dome-shaped roofs, almost standing still

- During the day, cumulus clouds increase but develop weakly vertically. Cloudy peaks have unclear contours and develop strongly around 15-16 hours. The peak of the cumulus cloud can be low, the cloudy feet spread out and fade away.
- The cumulus clouds are almost motionless; The amount of clouds does not increase over time or develop in a certain direction.
- Cumulus clouds form a wide area covering a part of the sky with a clear border.
- The daily variation of the temperature is evident. During the day, the air temperature increases moderately. At night the temperature drops due to the ground.
- In the hot season, there is a lot of fog in the evening, the night dew and salt fog form.
- In the coastal area, there is wind and earth - sea, windy mountain areas - strong blowing valley is a symptom of good weather.
- In the morning, dawn appears yellow light, sometimes dawn may start with a light red color, but then also turn yellow to show that the sun's rays meet relatively little steam on the way.

Weather phenomena tell the rain [6]

- Air pressure is constantly decreasing
- Cumulus clouds move quickly and are in a thin, parallel form. The faster clouds move, the faster the weather changes.
- The direction of movement of clouds in high above does not coincide with the wind direction at the bottom.
- At night, the wind is constantly blowing and strengthening.
- In the summer, the amount of clouds increases, the temperature decreases continuously.
- The sky at dawn and sunset is bright red without turning yellow, this phenomenon proves that the atmosphere contains a lot of water.

We find that when the pressure drops with low temperature and high humidity, it is a sign of rain. Because when the pressure drops will bring high air pressure to this place in the next 4 hours, these airflows will lead to clouds combined with high humidity and low temperature is a very visible phenomenon when rain.

2.1.2 Overview of forecasting [7]

The need to forecast how things work in the future has been around for centuries. The first predictions are forecasts of natural phenomena, social phenomena and phenomena of social life.

Initially, the forecast was based on the prediction of the forecaster, gradually forecasting was supported by the technology to help make the forecast more accurate.

Many predictive methods have been studied and released such as multiple regression analysis, Delphi method, Cross Impact Matrices, ARIMA (autoregressive integrated moving average), etc.

There are many ways to classify forecasts like:

- Based on the forecasting time, there are three types: long-term forecast, medium-term forecast and short-term forecast.
- Based on forecasting method: Forecast by expert method, forecast by regression equation, forecast based on time series, etc.

a. Concept of forecasting

Forecasting is a science and art that predicts things that will happen in the future.

Forecasting is scientific because it is based on the data collected from the past and based on the results of the analysis of factors affecting the forecast results.

The art of forecasting is based on practical experiences and judgment of experts to make predictions with the highest accuracy.

b. Purpose of forecasting

Make accurate, consistent decisions: Predictive analysis will provide detailed information about the forecasted audience from which strategic actions will be taken. Predictive analysis is carried out continuously and gives reliable results thanks to technical support. Decisions will be made consistently, fairly, not on human subjectivity.

Resolve tasks faster: Forecasting will answer complex questions and handle them with high accuracy in a short period of time. There are previous decisions that take hours or days, thanks to scientific support, only minutes or seconds.

Reduce costs due to risk reduction: Understanding the audience helps leaders to accurately assess risks and reduce losses.

c. Forecasting methods

The prediction methods are divided into two methods: qualitative methods and quantitative methods.

Qualitative method: Also known as an expert forecasting method. This method is used when data is not available or data is available but not enough for analysis and evaluation. Qualitative methods are often used when predictors are affected by factors that cannot be quantified. The principle of this method is to use the judgment of one or more experts in the relevant field. Therefore, to ensure the accuracy of the forecast, the subjective opinion of the forecaster must be excluded.

Quantitative methods: Use data from the past or collect data of current objects to forecast. With the quantitative method will give fast forecasting results and can measure the accuracy of the forecast. However, this method only applies to short and medium-term forecasts.

To predict high efficiency, people often combine both qualitative and quantitative methods.

2.2 Artificial Neural Network [8]

2.2.1 Introduction to Machine Learning

Machine learning is a field of artificial intelligence that involves researching and constructing techniques that allows systems to "learn" automatically from data and to solve specific problems. For example, systems can "learn" how to classify emails whether it is spam mail and automatically put them into the corresponding folder. Definition of machine learning is very close to statistical inference although there are differences between their terms.

Machine learning is highly relevant to statistics, since both areas related to data analysis. But unlike statistics, machine learning focuses on the complexity of algorithms in calculating performance. Many machine learning problems are classified as non-deterministic polynomial-time hardness problems, so part of machine learning is to study the development of approximate inference algorithms that can handle the problems.

Machine learning is now widely used including data tracing machines, medical diagnostics, stock market analysis, DNA sequencing, speech and writing recognition, facial recognition, automatic translation, and robot locomotion.

2.2.2 Introduction to Artificial Neural Networks

Artificial neural networks (ANNs) or connectionist systems are computing systems vaguely inspired by the biological neural networks that constitute the animal brain. Such systems "learn" to perform tasks by considering examples, generally without being programmed with any task-specific rules. In image recognition, for example, they might learn to identify images that contain cats by analyzing example images that have been manually labeled as "cat" or "no cat" and using the results to identify cats in other images. They do this without any prior knowledge about cats, e.g., that they have fur, tails, whiskers and cat-like faces. Instead, they automatically generate identifying characteristics from the learning material that they process.

An ANN is based on a collection of connected units or nodes called artificial neurons which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal from one artificial neuron to another.

In common ANN implementations, the signal at a connection between artificial neurons is a real number, and the output of each artificial neuron is computed by some non-linear function of the sum of its inputs. The connections between artificial neurons are called 'edges'. Artificial neurons and edges typically have a weight that adjusts as learning proceeds. The weight increases or decreases the strength of the signal at a connection. Artificial neurons may have a threshold such that the signal is only sent if the aggregate signal crosses that threshold. Typically, artificial neurons are aggregated into layers. Different layers may perform different kinds of transformations on their inputs. Signals travel from the first layer (the input layer) to the last layer (the output layer), possibly after traversing the layers multiple times.

The original goal of the ANN approach was to solve problems in the same way that a human brain would. However, over time, attention moved to performing specific tasks, leading to deviations from biology. ANNs have been used on a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games, and medical diagnosis.

2.2.3 Neural network models

The general neural network model takes the form of Figure 5

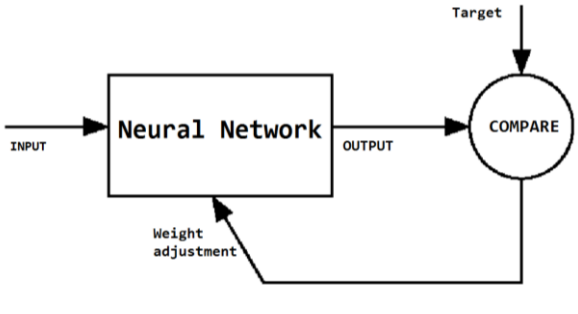


Figure 5: The general neural network model

a. Neural network basic models

i. Single-layer perceptron

The simplest kind of neural network is a single-layer perceptron network, which consists of a single layer of output nodes; the inputs are fed directly to the outputs via a series of weights. The sum of the products of the weights and the inputs is calculated in each node, and if the value is above some threshold (typically 0) the neuron fires and takes the activated value (typically 1); otherwise it takes the deactivated value (typically -1). Neurons with this kind of activation function are also called artificial neurons or linear threshold units.

A perceptron can be created by using any values for the activated and deactivated states as long as the threshold value lies between these states.

Perceptrons can be trained by a simple learning algorithm that is usually called the delta rule. It calculates the errors between calculated output and sample output data, then the result is used to create an adjustment to the weights, thus implementing a form of gradient descent.

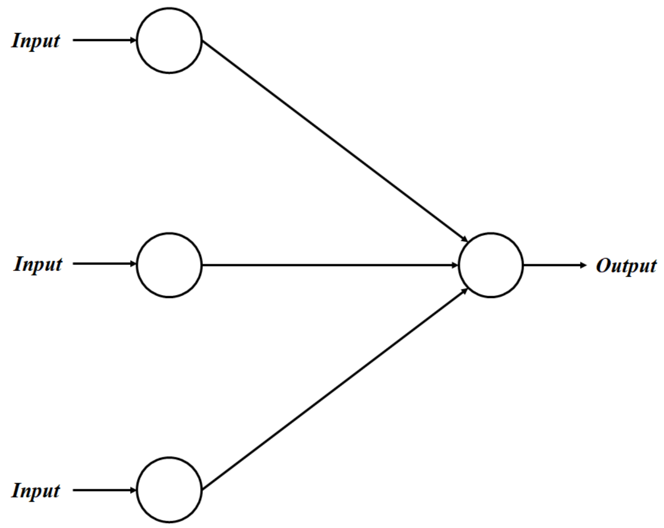


Figure 6: Single-layer perceptron model

ii. Multi-layer perceptron

This class of networks consists of multiple layers of computational units, usually interconnected in a feed-forward way. Each neuron in one layer has direct connections to the neurons of the subsequent layer. In many applications, the units of these networks apply a sigmoid function as an activation function.

The universal approximation theorem for neural networks states that every continuous function that maps intervals of real numbers to some output interval of real numbers can be approximated arbitrarily closely by a multi-layer perceptron with just one hidden layer.

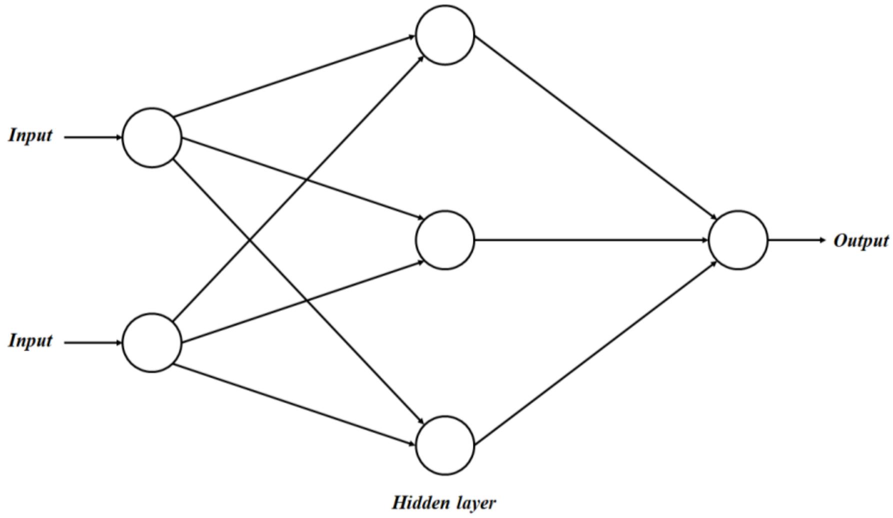


Figure 7: Multi-layer perceptron model

Multi-layer networks use a variety of learning techniques, the most popular being back-propagation. Here, the output values are compared with the correct answer to compute the value of some predefined error-function. By various techniques, the error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. After repeating this process for a sufficiently large number of training cycles, the network will usually converge to some state where the error of the calculations is small. In this case, one would say that the network has learned a certain target function. To adjust weights properly, one applies a general method for non-linear optimization that is called gradient descent. For this, the network calculates the derivative of the error function with respect to the network weights and changes the weights such that the error decreases. For this reason, back-propagation can only be applied to networks with differentiable activation functions.

b. Back-propagation neural network model

After two key issues with the computational machines that processed neural networks is discovered in 1969 by Minsky and Papert, a key trigger for renewed interest in neural networks and learning was Werbos's backpropagation algorithm. It effectively solved the exclusive-or problem by making the training of multi-layer networks feasible and efficient. Backpropagation distributed the error term back up through the layers, by modifying the weights at each node.

Backpropagation is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network. Backpropagation is shorthand for "the backward propagation of errors," since an error is computed at the output and distributed backward throughout the network’s layers. It is commonly used to train deep neural networks.

Backpropagation is a generalization of the delta rule to multi-layered feedforward networks, made possible by using the chain rule to iteratively compute gradients for each layer. It is closely related to the Gauss-Newton algorithm and is part of continuing research in neural backpropagation.

Backpropagation is a special case of a more general technique called automatic differentiation. In the context of learning, backpropagation is commonly used by the gradient descent optimization algorithm to adjust the weight of neurons by calculating the gradient of the loss function.

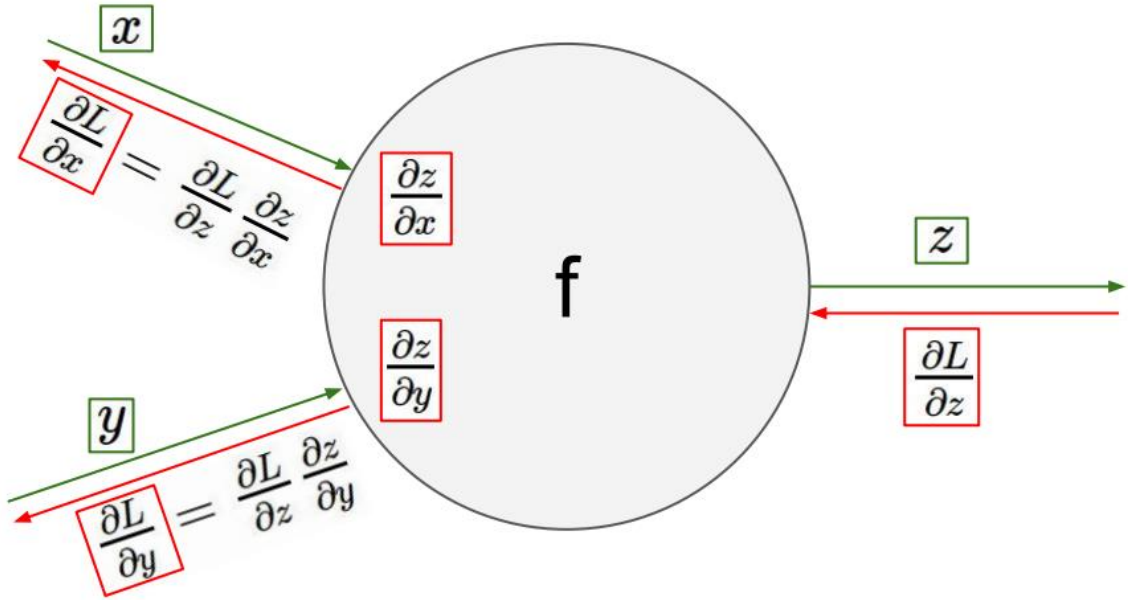


Figure 8: Demonstration of Back-propagation algorithm

Figure 8 shows the operation of the back-propagation algorithm at a neuron. The neuron has 2 input parameters x and y, the transfer function f and the result z. From z and the desired result, we calculate the reverse propagation error corresponding to the weighted adjustment number by the formula of the algorithm.

2.3 Components in the system

An overview of the components in the system of the topic is designed and connected to each other in Figure 9.

In Figure 9, the Environmental Monitoring Nodes are data collection devices for forecasting rainfall, which connects to the Gateways through LoRa connection. Gateways are a packet forwarding devices forwarding packets sent from the Environmental Monitoring Nodes to the Server by MQTT protocol via Wi-Fi. Server blocks play a role as the central processor of the entire system. The server stores all the weather data collected, manage the gateways and nodes. Specifically, MongoDB is chosen to be used as a database system. In addition, the modules in the Server are connected together via an internal socket.

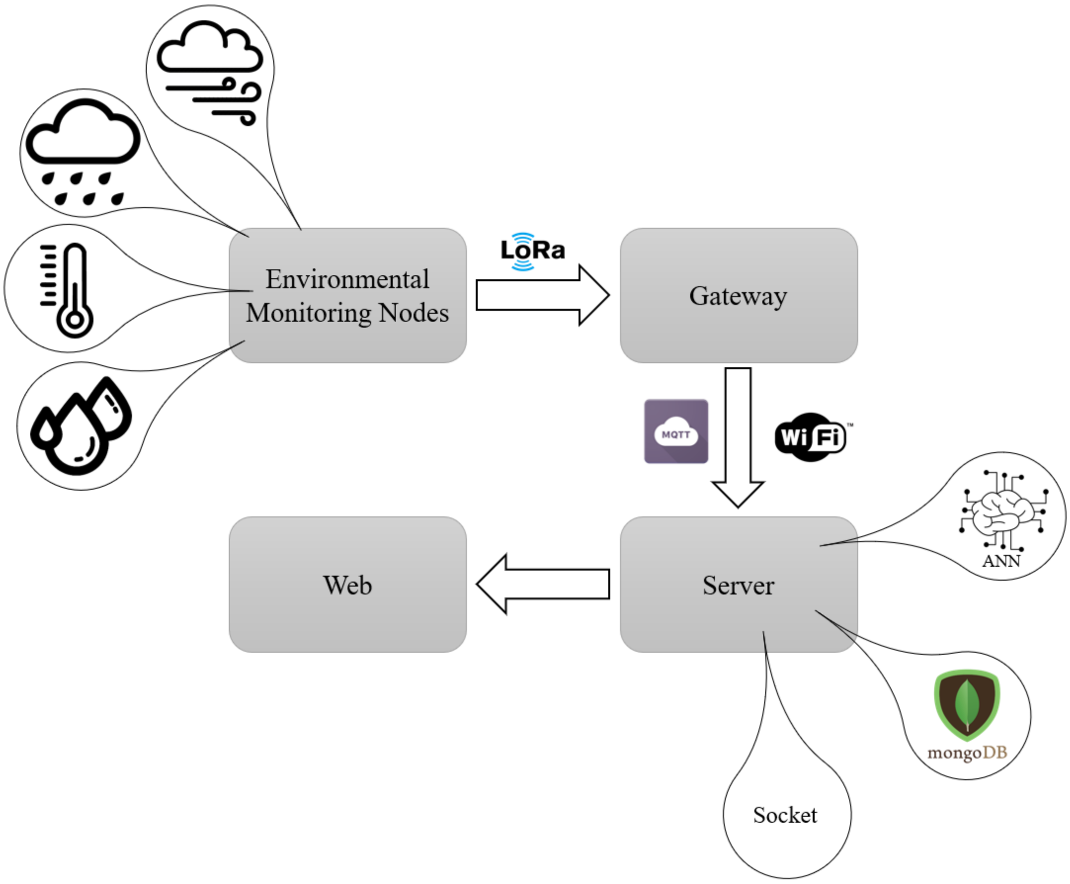


Figure 9: Overview of the components in the system

2.3.1 LoRa and LoRaWan [9, 10, 11]

LoRa, stands for Long Range Radio, is a wireless communication technology developed by Cycleo, and bought by Semtech in 2012. LoRa uses a modulation technique called Chirp Spread Spectrum. According to Semtech publication, this technique reduces the complication and accuracy requirement of receivers that need to achieve to decode data. In addition, LoRa does not require high transmit power to transmit to a receiver at a long distance cause the receiver can receive data even if received signal strength lower than ambient noise.

The technology is presented in two parts — Lora, the physical layer and LoRaWan, the upper layers. LoRaWan is a protocol based on LoRa developed by LoRa Alliance. It's used in the industrial, scientific and medical (ISM) radio bands. A LoRaWan network structure usually is a star-of-stars model. Gateways are bridges forwarding messages between end nodes and server at the backend. Gateways connect to the internet via normal IP connections. Meanwhile, end nodes wirelessly single-hop connect to one or many gateways.

Figure 10: Working bandwidths of LoRa for regions of the world [11]

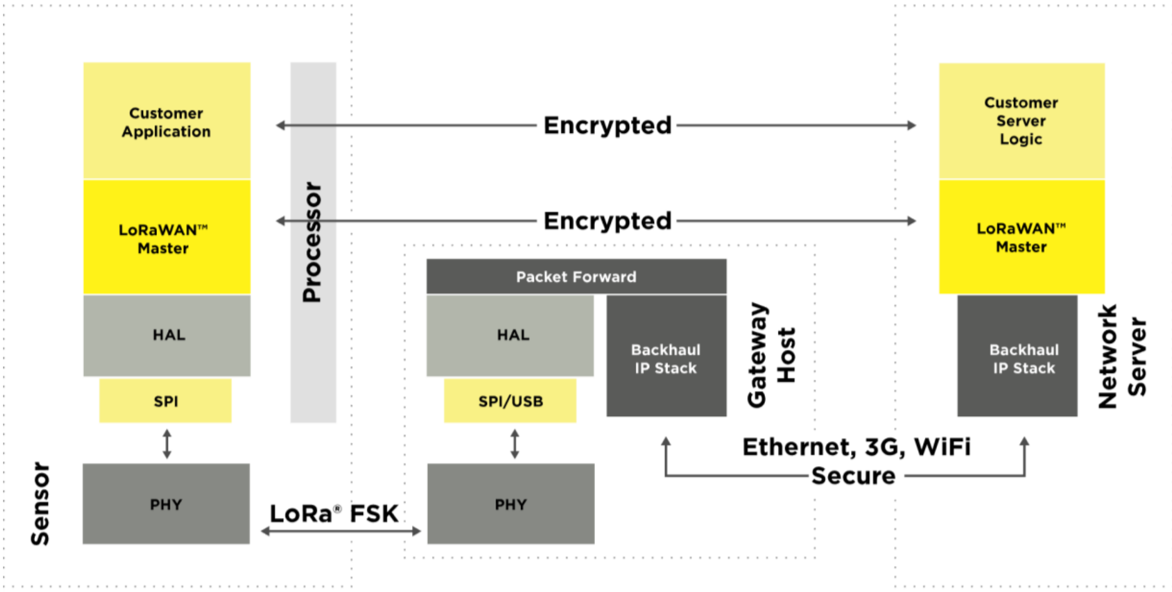


Figure 11: A LoRaWan system architecture [10]

A single LoRaWan Gateway can cover up to 10Km. With the benefits of LoRaWan's distance, The Things Network covers Amsterdam with just 10 gateways and costs $1200.

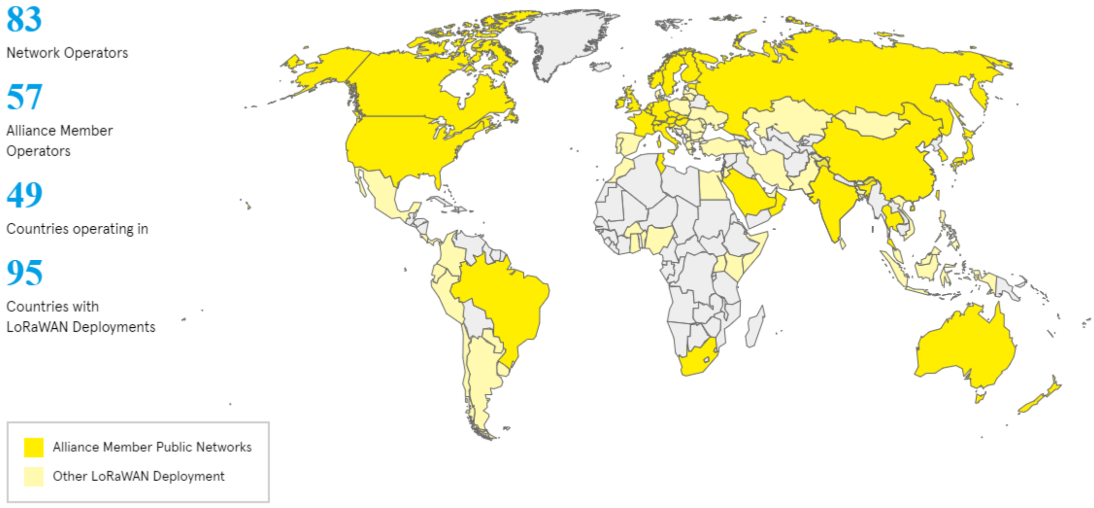


Figure 12: LoRaWan Cover Area in the world [11]

2.3.2 MQTT [12]

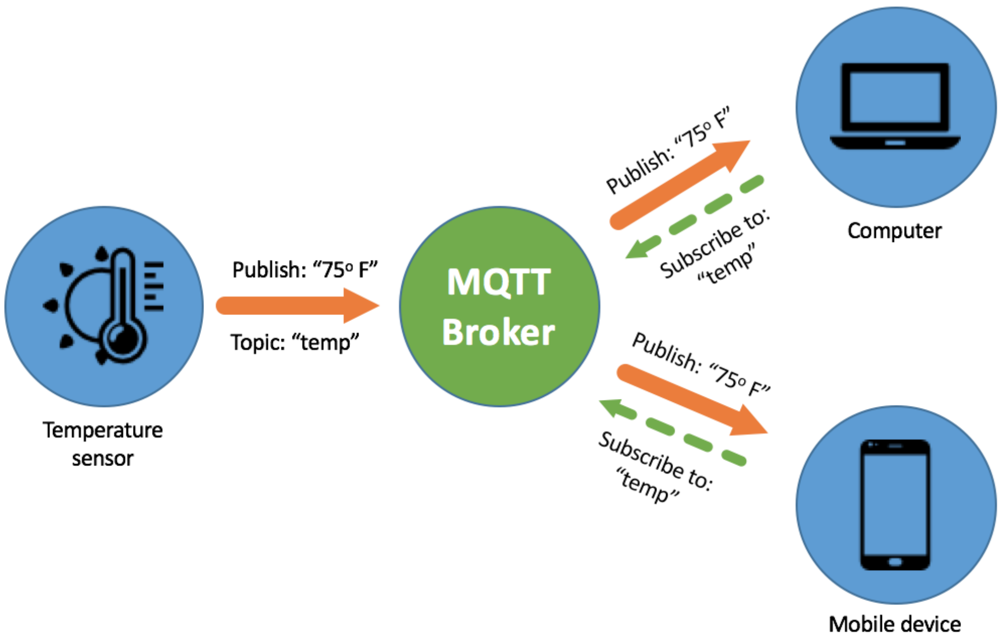


Figure 13: Structure of a MQTT system [12]

MQTT was developed by IBM and Eurotech, the latest version is MQTT 3.1.1. MQTT (Message Queuing Telemetry Transport) is a publish / subscribe protocol commonly used for Internet of Things devices with low bandwidth, high reliability and the ability to be used in unstable networks. It is based on a Broker and is designed to be open and non-specific to any application, very simple and easy to integrate. MQTT is suitable for M2M (Mobile to Mobile) applications, WSN (Wireless Sensor Networks) or IoT (Internet of Things).

There are 3 levels of QoS (Qualities of service) in MQTT. Figure 14 gives a demonstration of these qualities of service:

- QoS 0 – Almost once: The broker/client sends the data once, only TCP/IP protocol confirm the successful receiving
- QoS 1 – At least once: The broker/client sends at least one successful receiving packet to the sender
- QoS 2 - Exactly once: The broker/client sends only one packet to the sender to confirm that the receiving process is successful

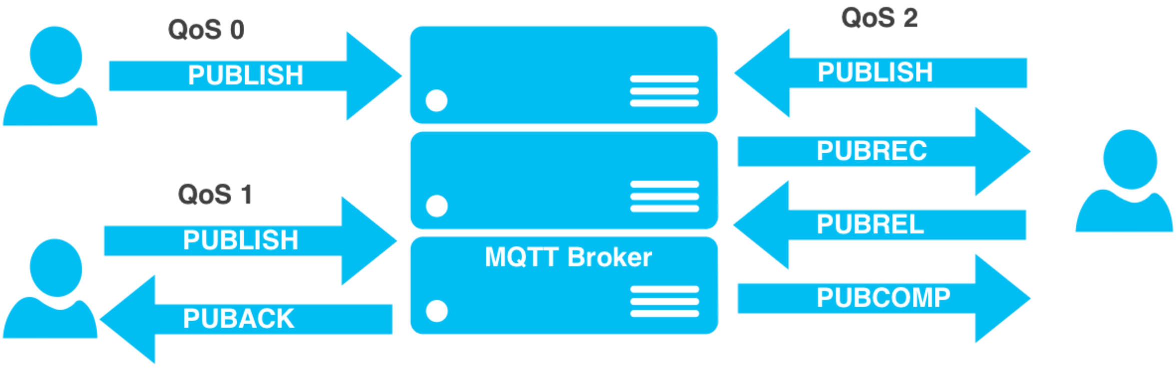


Figure 14: A demonstration of MQTT QoS [12]

2.3.3 Unix Socket [13]

Sockets enable communication between two processes in the same computer or different computer. It's a method to communicate with other processes via standard Unix file descriptors. In the Unix environment, every I/O operation is done by reading or writing to a file descriptor. A file descriptor is an integer associated with an open file such as text files, network connections, etc.

Sockets were first introduced in 2.1BSD and edited to current form in 4.2BSD. Socket feature is available to most popular UNIX system.

A Unix Socket is usually used in a server-client framework. Most protocols like FTP, SMTP and POP3 use socket to establish a connection between server and client to transmit data.

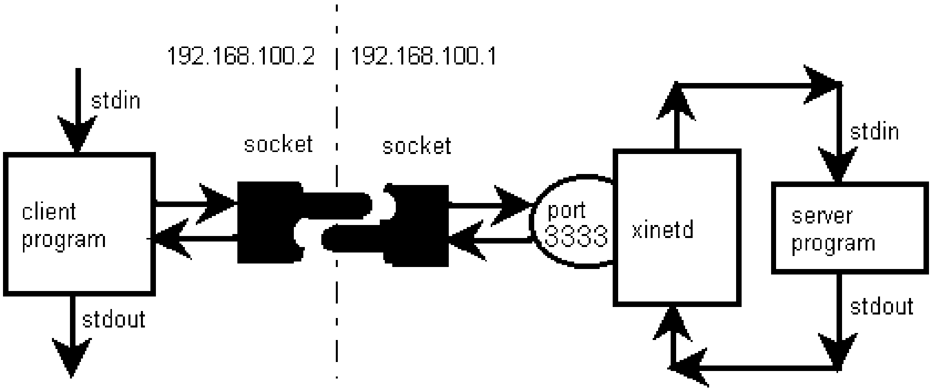


Figure 15: Sample flow chart of a server-client application

There are four types of sockets available to the users:

- Stream Sockets
- Datagram Sockets
- Raw Sockets
- Sequenced Packet Sockets

2.3.4 MongoDB [14]

MongoDB is a cross-platform document-oriented database. It is higher performance and easier to scale than RDBMS. Documents of MongoDB have the same structure as JSON. MongoDB is first introduced by MongoDB Inc in February 2009. After that, MongoDB quickly developed into one of the most popular NoSQL databases.

```
{
    "_id" : ObjectId("5b56deba65783d1c6832f7e4"),
    "value" : "40b7169e-3b35-4944-a8c4-5f1c2363c5a01532419763440",
    "topic" : "$SYS/ipTStsN/disconnect/clients",
    "options" : {
        "qos" : null,
        "messageId" : "iLgNhl1"
    }
}
```

Figure 16: A sample JSON-like document in MongoDB

The current version of MongoDB is 4.1.1 published at Git Repository: github.com/mongodb/mongo.

RDBMS	MongoDB
Database	Database
Table	Collection
Index	Index
Row	Document
Column	Field
Join	Embedding & Linking

Figure 17: Relational concepts in RDBMS [14]

Main features of MongoDB

- Schema less
- High availability by cluster
- Structure of a single object is clear
- No complex joins
- Deep query-ability
- Tuning
- Ease of scale-out
- Uses internal memory for storing the (windowed) working set,

enabling faster access of data

Use case

- Big Data
- Content Management

2.3.5 Sensors and modules

Weather Station SKU:SEN0186 [15]

Equipped with a rain bucket and wind accelerometer, the SKU:SEN0186 from DFRobot is a popular solution to many weather monitoring applications at a reasonable price.

Figure 18: A real image of SKU:SEN0186 [15]

Specification

- Operating voltage: 5V
- Temperature range: -40~80℃
- Humidity range: 0~99%
- Package Dimension: 20*18*30 CM
- Weight: 4480g

Application

- Weather station

- Weather monitor

Data interface

Serial: 9600bps with 1s interval

Format of Data Output

35 bytes per second, including the end CR/LF

For example: c000s000g000t086r000p000h53b10020\r\n

c000： Air direction, 0 degree

s000： Air speed (1 minute), 0 miles per hour

g000： Air speed (5 minutes), 0 miles per hour

t086： Temperature, 86 Fahrenheit Degree

r000： Rainfall (1 hour), 0 inches

p000： Rainfall (24 hours), 0 inches

h53： Humidity, 53%

b10020： Atmosphere, 1002 hpa

Temperature/Humidity Sensor SHT10 [16]

SHT10 is designed by Sensirion The Sensor Company with a 14-bit-analog-to-digital converter and a serial interface circuit. It is a low-cost series sensor comparing to other SHT1x series.

Specification

- Series: SHT10
- Type Humidity Accuracy %: ±3
- Type Temperature Accuracy °C: ±0.4
- Supply Voltage Range V: 2.4 to 5.5
- Interface: Digital SBus
- Package Size: 7.5mm x 4.9mm x 2.6mm

Figure 19: Images of SHT10 with metal stainless steel protective cover [16]

The version used in this thesis is modified with protective case protecting the sensor from physical environmental factors like water, dust, etc.

Barometric Pressure Sensor BMP180 [17]

BMP180 is allow-cost sensing solution for measuring barometric pressure and temperature produced by Bosch. Because of pressure changes with altitude, this sensor can also be used as an altimeter.

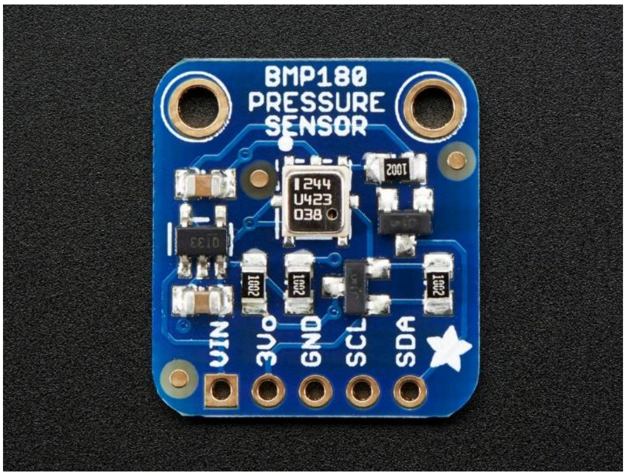


Figure 20: An image of BMP180 [17]

Specifications

- Operating voltage: 1.8 ~ 3.6V
- Operating Current: 0.5uA at 1Hz

- Interface: I2C
- Pull-up resistors integrated on the I2C pins
- Max interface speed: 3.5MHz
- Low error: 0.03hPa (25cm)
- Integrated calibrate module available
- Pressure Measuring range: 300hPa ~ 1100hPa (+9000m to -500m)
- Weight: 1.18g
- Package size: 21mm x 18mm

Ultra-Long-Range Transceiver RFM95 [18]

The LoRa transceiver module RFM95 produced by HopeRF provides ultra-long-range communication with minimizing current consumption. Using the Chirp Spread Spectrum modulation technique, the RFM95 can reduce the sensitivity to ~ -148dBm, roughly equal to the ambient noise level. Combined with the 20dBm Power Amplifier, the communication distance can reach up to 15km.

Figure 21: An image of RFM95 [18]

Specifications

- LoRa™ Modem
- 168 dB maximum link budget
- +20 dBm - 100 mW constant RF output vs V supply

- +14 dBm high efficiency PA
- Programmable bit rates up to 300 kbps
- High sensitivity: down to -148 dBm
- Bullet-proof front end: IIP3 = -125 dBm
- Excellent blocking immunity
- Low RX current of 103 mA, 200 mA register retention
- Fully integrated synthesizer with a resolution of 61 Hz
- FSK, GFSK, MSK, GMSK, LoRa™ and OOK modulation
- Built-in bit synchronizer for clock recovery
- Preamble detection
- 127 dB Dynamic Range RSSI
- Automatic RF Sense and CAD with ultra-fast AFC
- Packet engine up to 256 bytes with CRC
- Built-in temperature sensor and low battery indicator
- Package Size： 16*16mm

Applications

- Automated Meter Reading
- Home and Building Automation
- Wireless Alarm and Security Systems
- Industrial Monitoring and Control
- Long range Irrigation Systems

Wi-Fi Module ESP8266-V12 [19]

The ESP8266-V12 is a low-cost Wi-Fi module which is highly rated for Internet and Wi-Fi applications as well as transmission applications. The ESP8266 is an integrated chip designed for the needs of the Internet of Things (IoT). It provides a complete and closed Wi-Fi network solution that allows it to host applications or to reduce the load on all Wi-Fi network connections from the controller. The ESP8266 has powerful processing and storage capabilities that allow it to be integrated with other sensors, microcontrollers and application devices via GPIOs.

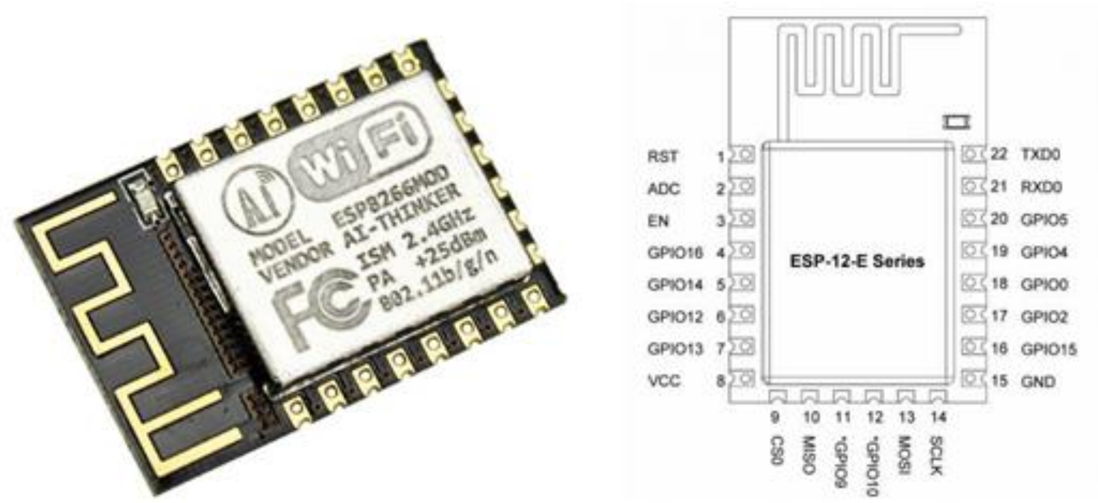


Figure 22: Images of ESP8266-V12 [19]

Specifications

- Producer: Espressif
- Certification: Wi-Fi Alliance
- Protocols: 802.11 b/g/n
- Frequency Range: 2.4G ~ 2.5G
- CPU: Tensilica L106 32-bit processor
- Peripheral Interface: UART/SDIO/SPI/I2C
- Operating Voltage 2.5V ~ 3.6V
- Operating Current: ~80 mA
- Operating Temperature Range: –40°C ~ 125°C

2.4 IDEs, software, and tools

2.4.1 Git and GitHub

Git is a version control system developed by Linus Torvalds in 2005, with the purpose of developing Linux kernel at the beginning. Nowadays, Git becomes the most popular version control system. Git is an open source application with the GPL2 license.

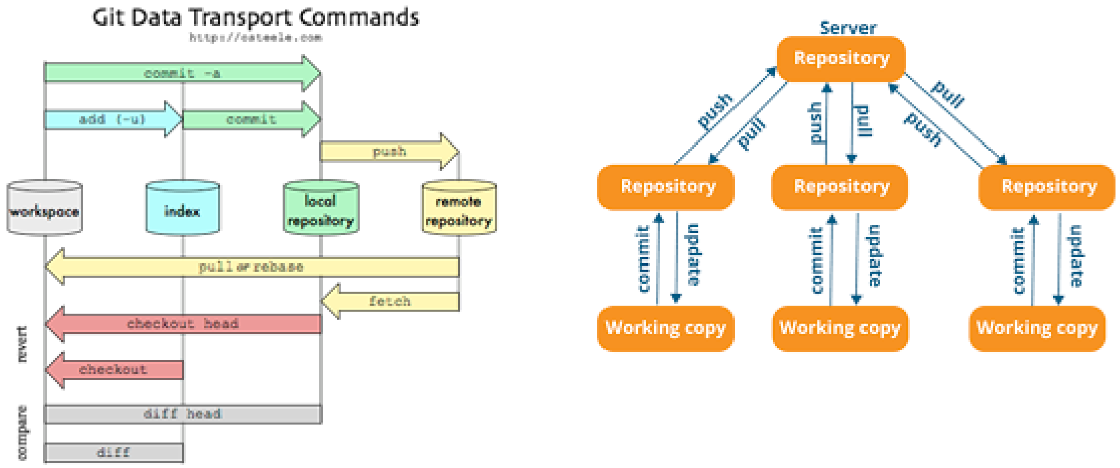


Figure 23: Stages and operations of Git

Designed by Linus Torvalds

Developed by Junio Hamano and others

First publish in April 7th, 2005

Repository https://git.kernel.org/pub/scm/git/git.git/

Platform Linux, Windows, OS X

Homepage git-scm.com

GitHub is a service that provides web-based Git source code repositories for software development projects. GitHub offers both paid and free versions for accounts. Open source projects will be offered free storage. As of June 2018, GitHub has more than 40.1 million users, making it the largest source code server in the world. GitHub has become a trend in the open source development community. Even many developers have begun to consider Git as a substitute for resumes and some employers require applicants to provide a link to the GitHub account to evaluate candidates. Formally acquired by Microsoft on June 4th, 2018 for $ 7.5 billion.

Figure 24: GitHub’s logo and mascot

Founded in February 8th, 2008

Headquarters in San Francisco, California, U.S

Founders: Tom Preston-Werner, Chris Wanstrath, PJ Hyett

Our Git repository is at https://github.com/nguyenmanhthao996tn/ktln

The repository has 3 main folders:

- **Documents**: Contains documents like related papers that relate to this thesis, report, testing results, cost, etc.
- **PCBs**: Contains all hardware designs.
- **SourceCode**: Contains source code of server and firmware of other components

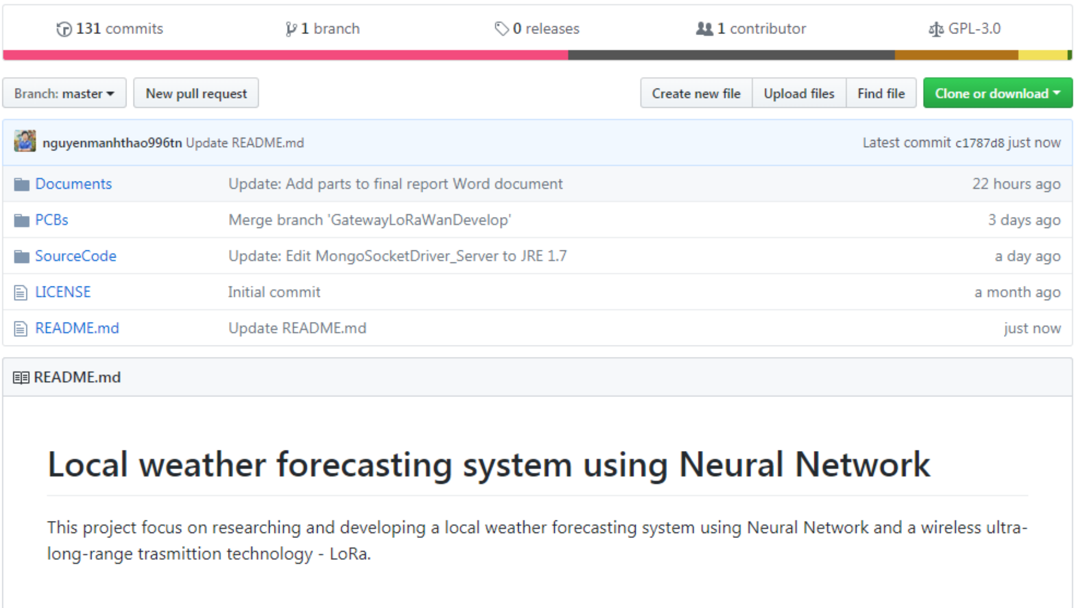


Figure 25: Repository of this thesis

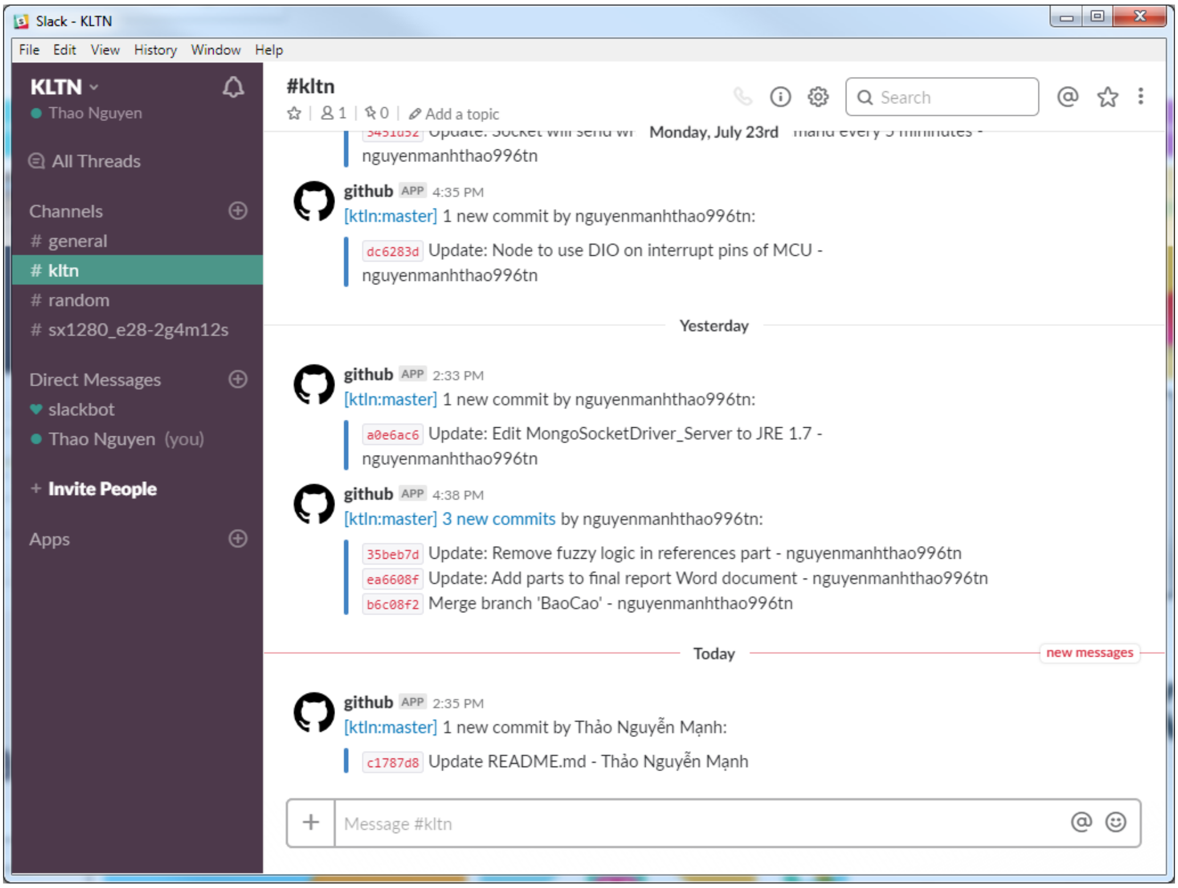


Figure 26: Our Slack channel with Git notification add-on integrated

2.4.2 Slack

Slack is a chat platform or a set of proprietary team collaboration tools and services. First introduced in August 2013, Slack is developed by Slack Technologies and its original author is Stewart Butterfield. Slack's users can be able to join conversations on different channels. Moreover, Slack provides many features, integration tools, and plugins. Typically, Git notifications are integrated through GitHub's Hook feature. Besides, it's free for everyone.

2.4.3 VirtualBox

VirtualBox is a free, open-source powerful x86 and AMD64/Intel64 virtualization software. It supports on many platforms such as Windows, Linux, MacOS. VirtualBox is compatible with most 32-bit and 64-bit operating systems but not require additional virtualized hardware. In additional, VirtualBox provides a powerful Guest Additions mode that helps user can interact to the virtual machine much easier.

Original author: Innotek GmbH

Developer: Oracle Corporation

Initial release 17 January 2007

Repository: https://www.virtualbox.org/browser/vbox/trunk

Written in: C, C++, x86 Assembly

Operating system: Windows, macOS, Linux

License: Base Package: GNU General Public License version 2

2.4.4 Visual Studio Code

Visual Studio Code is a combination of a text editor and supporting tools that developers need such as an integrated terminal, Git client, plugins, code formatting feature, etc.

Visual Studio Code is a free, open source, cross-platform. Developed by Microsoft and first published in April 29th, 2015. It's repository: github.com/Microsoft/vscode

Chapter 3. **SYSTEM ANALYSIS AND DESIGN**

3.1 System overview

The entire system design is shown in the following Figure 27:

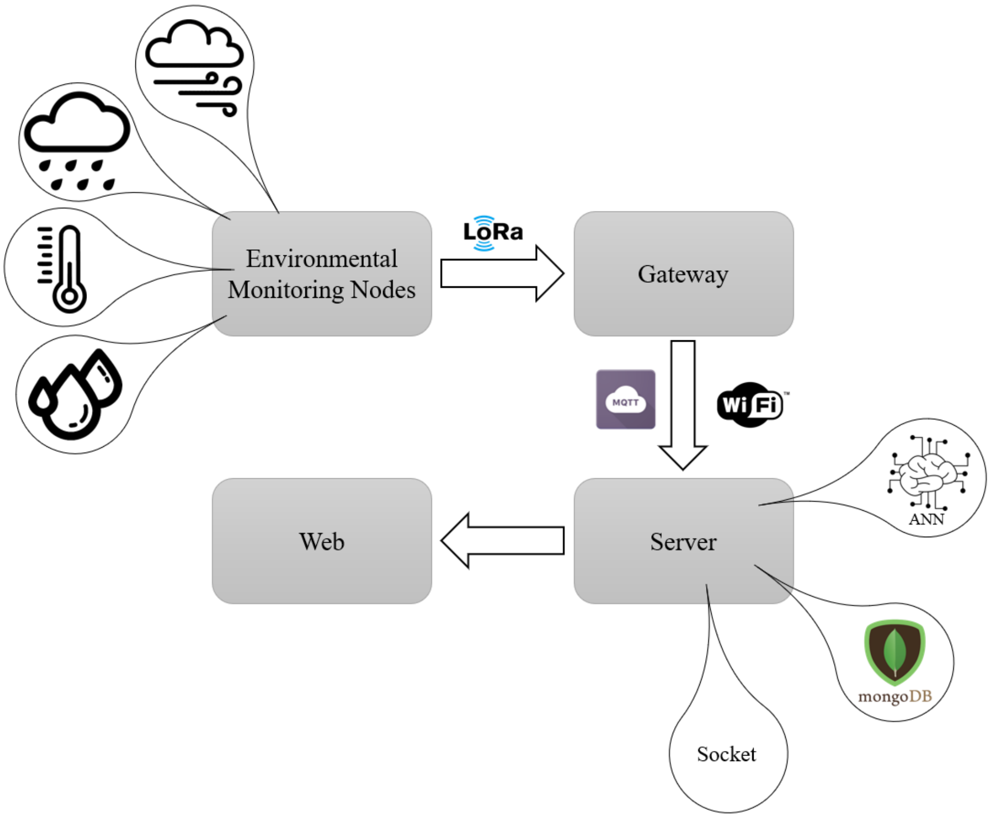


Figure 27: System design chart

Environmental Monitoring Nodes are responsible for collecting environmental data at the place where the weather needs to be forecasted and sending this data to the server through the gateway every 5 minutes. Data collected include wind direction, wind speed, temperature, humidity, pressure. The output of this unit includes air direction, airspeed in 1 minute, airspeed in 5 minutes, temperature, rainfall in 1 hour, rainfall in 24 hours, humidity, and atmosphere.

Gateways are responsible for forwarding packets from the Environmental Monitoring Nodes to the server for processing via the Internet, Managing the Nodes in the LoRa covered area, confirming the packet and requesting nodes to resend the packet when receiving packets fail.

The server receives packets from Gateways, processes and saves the weather data to the database. After that, this data is used to predict rainfall in 4 hours in advance for each individual area. The prediction is triggered once per hour and the learning process is triggered once a day. In addition, the server provides API to send data to users through a variety of methods such as Website.

User Interface provides environmental data and predicted results via two methods, which are LCD integrated on the Gateway and Website.

3.2 Hardware design

3.2.1 Environmental Monitoring Node

Environmental Monitoring Nodes are design based on the following criteria: Power saving, easy to install, durable for outdoor operating in the rain weather.

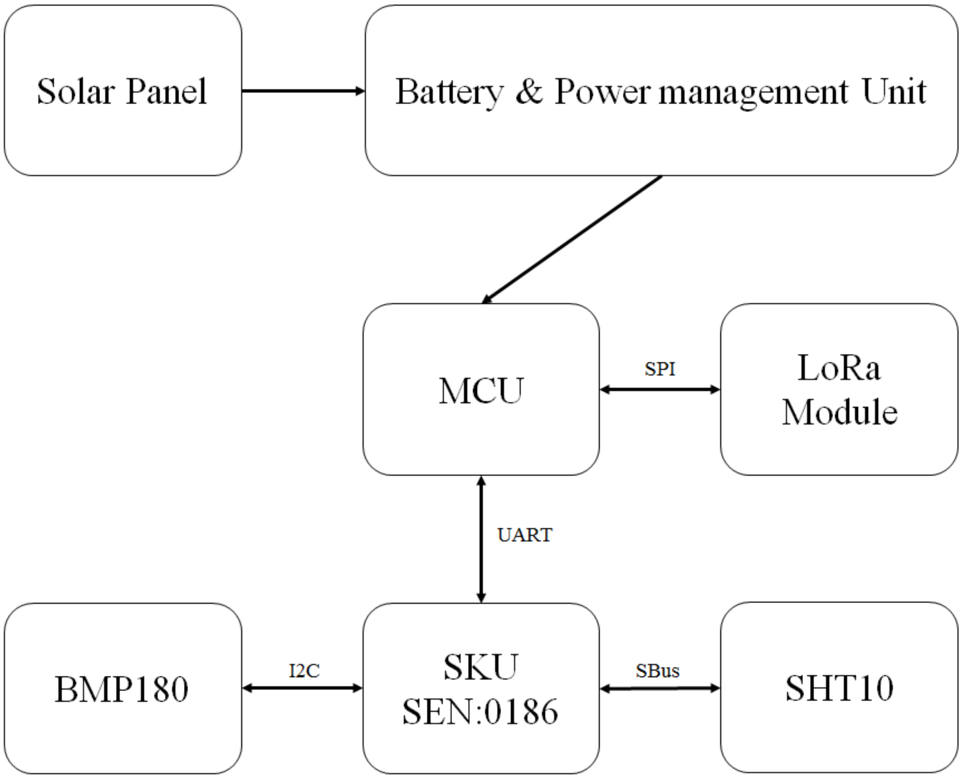


Figure 28: Hardware design of Node

Figure 29: An image of the Environmental Monitoring Node

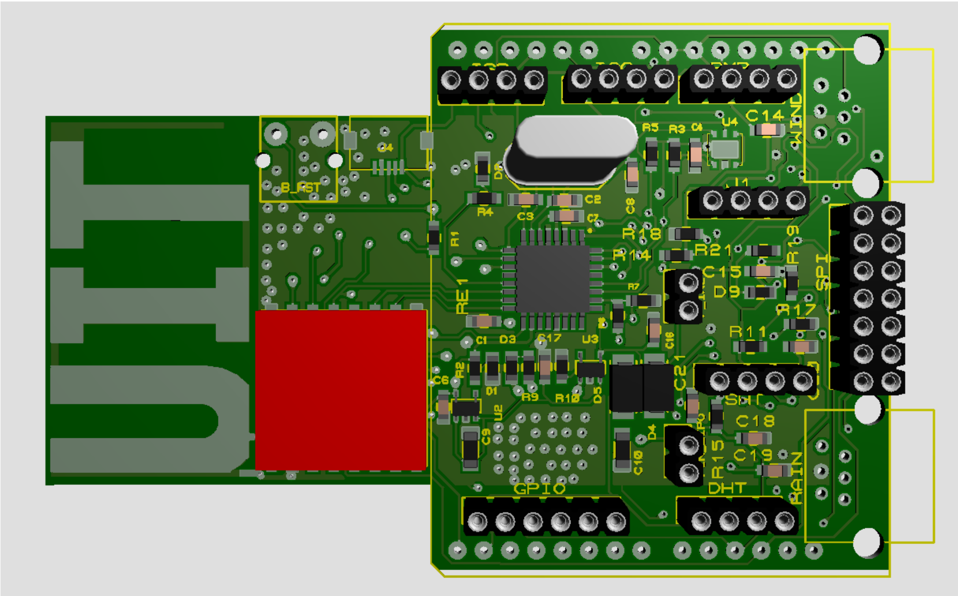


Figure 30: 3D simulation of the Environmental Monitoring Node circuit

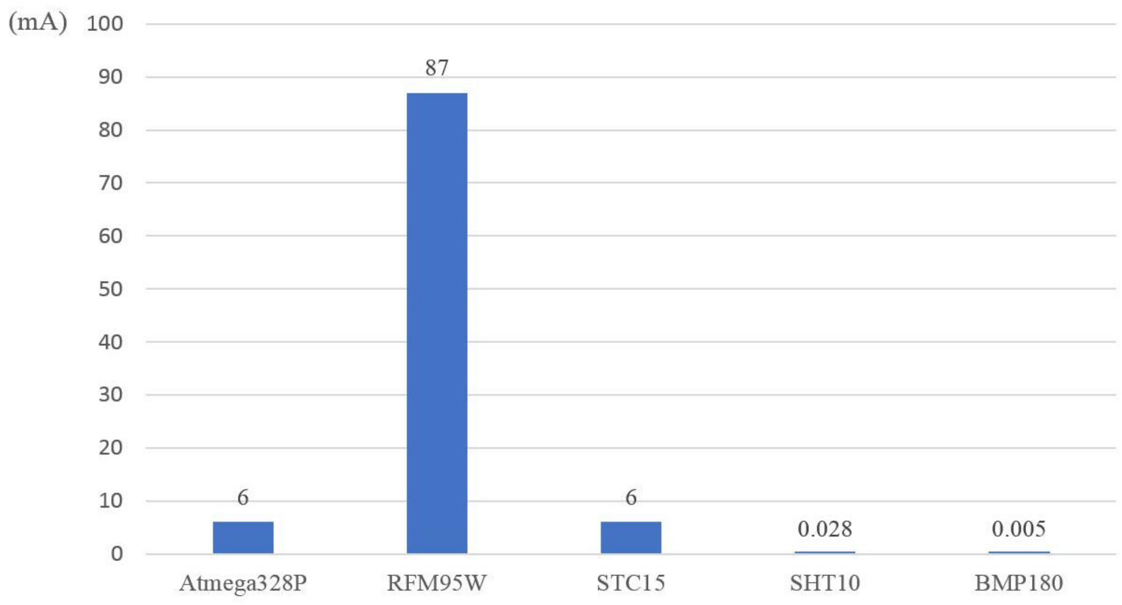


Figure 31: Device operating currents in normal mode

For more detail:

- Atmega328P draws ~6mA at Active mode 3v3 16MHz
- RFM95W draws 87mA at +17 dBm
- STC15 draws 6mA
- SHT10 draws 28µA at 1 measurement of 12bit resolution per second
- BMP180 draws 5µA at 1 sample / sec. in standard mode

The Figure 31 above describes the operating current of Weather Monitoring Node components. There is a lot of energy consumed in transfer operations. For that reason, we decide to use sleep mode in the LoRa module and MCU (Atmega328P) to save power.

For MCU (Atmega328P), we disable unused modules: ADC, SPI, Timer0, Timer2, TWI and put it to Idle mode, which is not disable USART module. Operating current that we can save when disable unused modules above:

- ATMega328P draws ~1.2mA in Idle mode compare to 6mA in Active mode

- ADC draws 295.38µA
- SPI draws 186.5µA
- Timer0 draws 61.13µA
- Timer0 draws 224.25µA
- TWI draws 199.25µA

Total operating current of MCU on node:

- Sleep mode: 1200 - (295.38 + 186.50 + 61.13 + 224.25 + 199.25) = 233.49µA = 0.23349mA
- Active mode: 6000 - (295.38 + 186.50 + 61.13 + 224.25 + 199.25) = 5033.49µA = 5.03349mA

In addition, MCU needs to active to get 35 bytes data from STC15 with baud rate 9600 every 5s.

Table 1: Operating current of MCU on Weather Monitoring Nodes

Time (s)	4.97	0.03
Power (mA)	0.23349	5.03349

For RFM95W, it draws 0.2uA in sleep mode and needs to active every 5 minutes to send data.

Table 2: Operating current of RFM95W on Weather Monitoring Nodes

Time (s)	4.97	0.03
Power (mA)	0.23349	5.03349

In conclusion, MCU draws 262.29µA and RFM95W draws 233.97µA in average. Therefore, the entire node draws 262.29µA + 233.97µA + 6mA + 28µA + 5µA = 6529.26µA = 6.52926mA.

The battery we use for this system is a pair of Panasonics 18650 2800mAh 3.7V. Total capacity that can supply the system is 2800x2=5600mA. So, the lifetime of the system with full battery is 5600/6.52926 = 857.68 hours = 35.74 days according to theory.

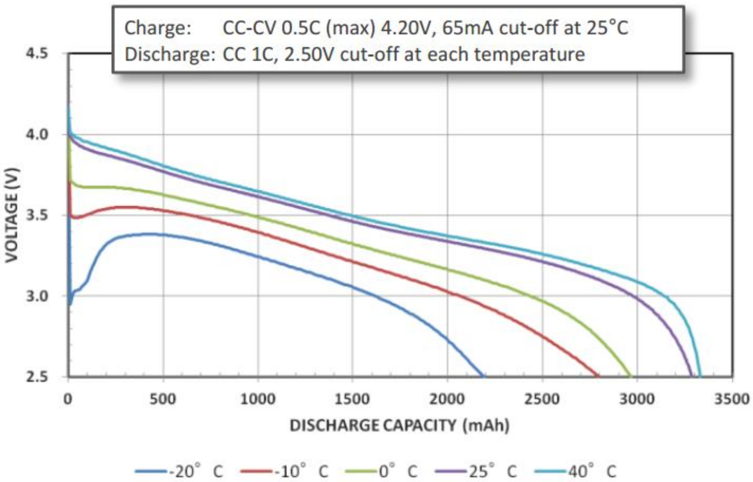


Figure 32: Battery discharge capacity versus voltage

3.2.2 Gateway

Gateway design includes 2 different versions: simple version using Wi-Fi connection and a version using Raspberry Pi. Each version has different pros and cons.

For the simple version using Wi-Fi, the biggest advantage is low-cost, simple and small design. It includes only 2 components which are a LoRa RFM95W module and a Wi-Fi ESP8266 module. Total price is under 10$ for each gateway of this kind. However, it only works with the Wi-Fi connection.

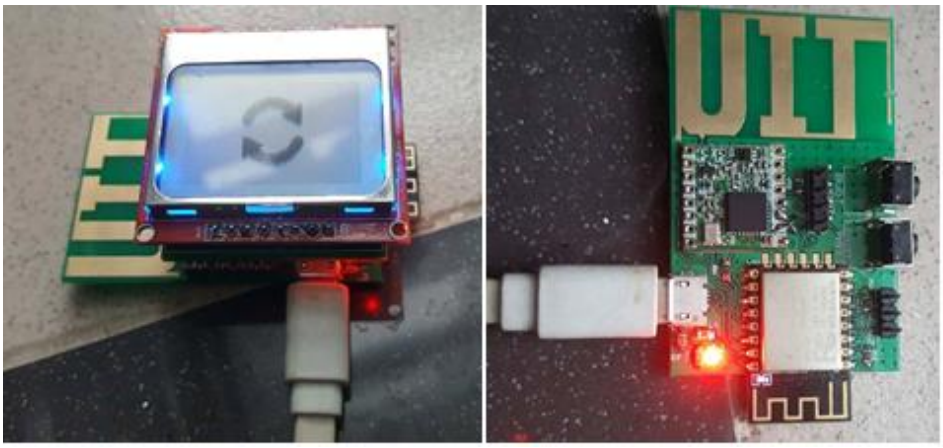


Figure 33: Gateway using ESP8266 with LCD

For Raspberry Pi version, the operating area radius depends on the antenna and usually higher than ESP8266 version. The antenna is connected via IPEX or SMA header. In addition, each gateway uses two LoRa RFM95W modules, so it can manage more node in an area than ESP8266 version even in multi-channel. Besides, the Internet connection does not limit to Wi-Fi, it can use Ethernet as well. The application expand capability in this version is also a very good cause of the OS we use is Raspbian, which is a distribution of Linux Debian base.

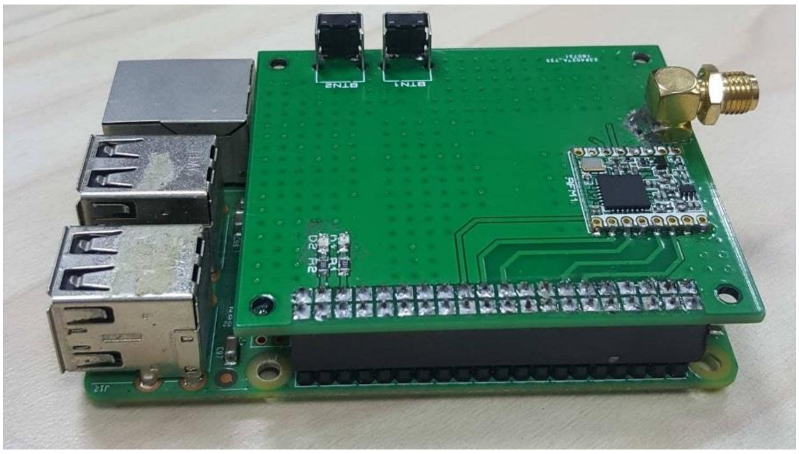


Figure 34: Gateway using Raspberry Pi 3

Both version support LoRaWan standard and an external LCD is available upon request. The currently supported LCD is LCD Graphic Nokia5110 with the resolution is 84x48 pixel.

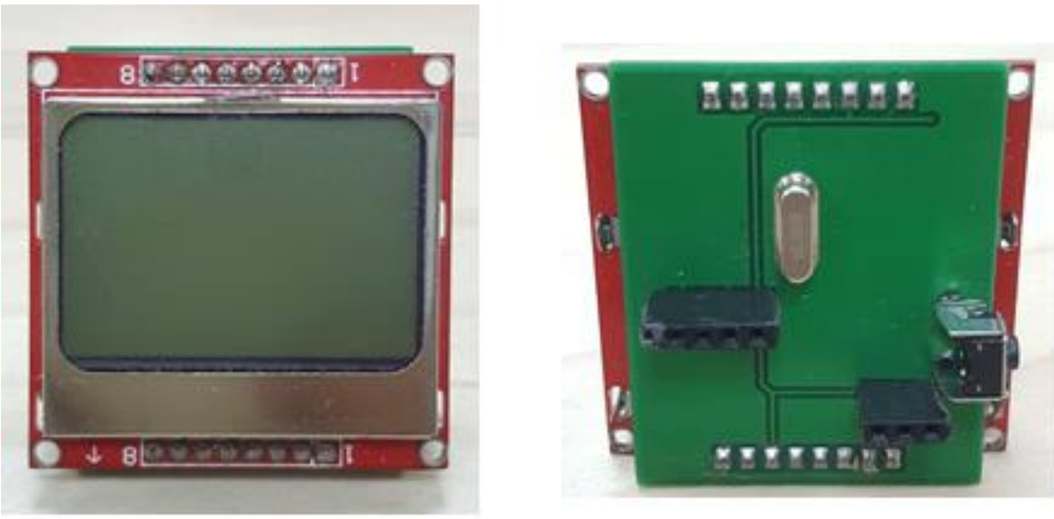


Figure 35: Support LCD Graphic Nokia5110

3.3 Software design

The applications and services running on the server are designed to handle various tasks such as database management, Gateway connection management, providing API for the website to retrieve data, trigger learning and predicting process of the forecasting model. These tasks are divided into several units, including the database management unit, Gateway connection management unit, machine learning model management unit, website management unit. These units link and interact with each other as shown in Figure 36:

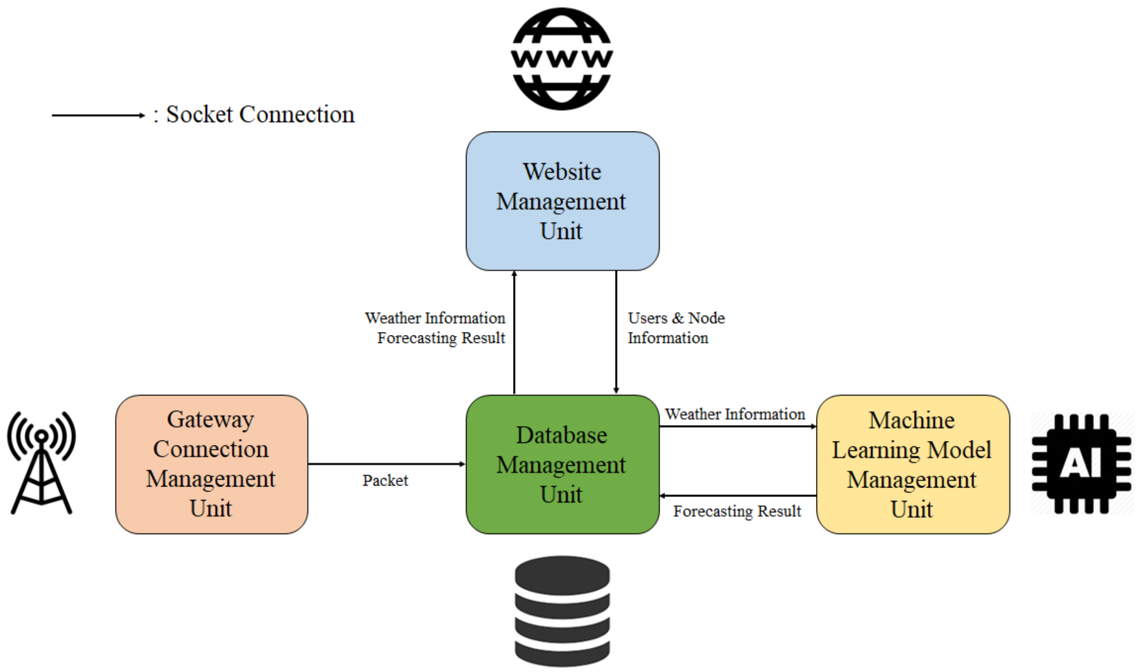


Figure 36: Server block diagram

Database Management Unit

Designed for the purpose of big data processing, this unit used MongoDB as a data storage platform. We installed the database application on the server and designed the data access management application on this database to avoid overloading the database when giving direct database access from other units. Data access management application is written in Java and uses MongoDB library provided by MongoDB Inc. This application allows other blocks to indirectly interact with the data stored through the internal socket connection at port 5001. For example, a request to write a predicted result from the machine learning model management unit has the following structure:

```
{
   "code":2,
   "nodeId":"Weather_Node_1",
   "date":"2018-11-10 13:15",
   "rain":1,
   "AmountOfRain":20
}
```

The structure of this unit is also specially designed for the request handler not to cause a failure or a slow response when too many requests are sent at the same time. The organization of this algorithm can be described as in Figure 37.

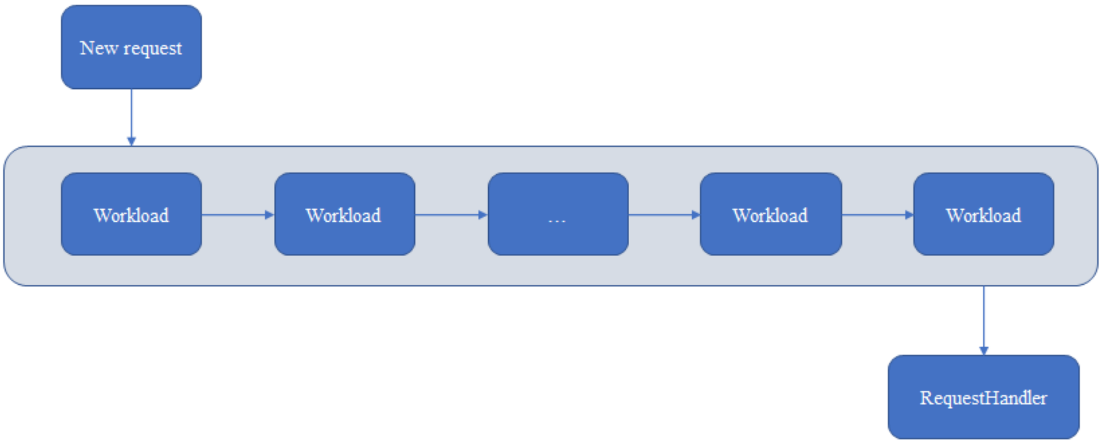


Figure 37: Flow of requests in database management unit

In this structure, Workload is the decoded task of the request string received from the socket and queued to be processed, RequestHandler is the object to handle Workload saved in the queue. Currently, only 4 RequestHandlers is created in the whole unit running on 4 independent and parallel processes. Creating more RequestHandler will consume more system resources. However, it will reduce the waiting time of requests when multiple requests come at the same time. Choosing 4 RequestHandler in this unit is tested via the delay of the request by "Stress Test" method.

Gateway Connection Management Unit

This block has the main function of managing connections and packets transferred from gateways. It is designed on the basis of information sharing and expandable ability. The criteria which we want to focus on is the expansion of the machine learning model, which will use a lot of input data collected from many monitoring devices around areas instead of one monitoring devices only.

The Gateway Connection Management Unit is developed by NodeJS and connect to the Database Management Unit via socket connection to write received packets. Gateways connect to this unit via the MQTT protocol, a protocol created to shares information between devices like a social network. Packets received from monitoring devices will be forwarded to The Gateway Connection Management Unit from these gateways, where the packets will be processed again to make sure that there are no errors in the packet during transmission. Finally, the connection management block from the Gateway will create a request to write the received data into the database, used for predicting the rain as well as future developments.

Machine Learning Model Management Unit

The main function of this unit is to manage the machine learning model, schedule and control the learning process of the model, trigger the rain prediction. When designing this model, we examined and gave a table comparing the advantages and disadvantages of machine learning models as mentioned in 2.1.2. Based on this comparison, we have decided to use the Perceptron and Back-propagation Multi-layer algorithm model to solve the rain prediction problem and built this model in C/C++ language. Details will be described later in section 3.4.

Website Management Unit

The website management unit is responsible for handling web-related tasks like a web server, specifically taking environmental data, predicting results, information about monitoring devices and then displaying it on the web. Developed in NodeJS and like other blocks, this block binds to the database management unit via the socket protocol to retrieve the necessary data.

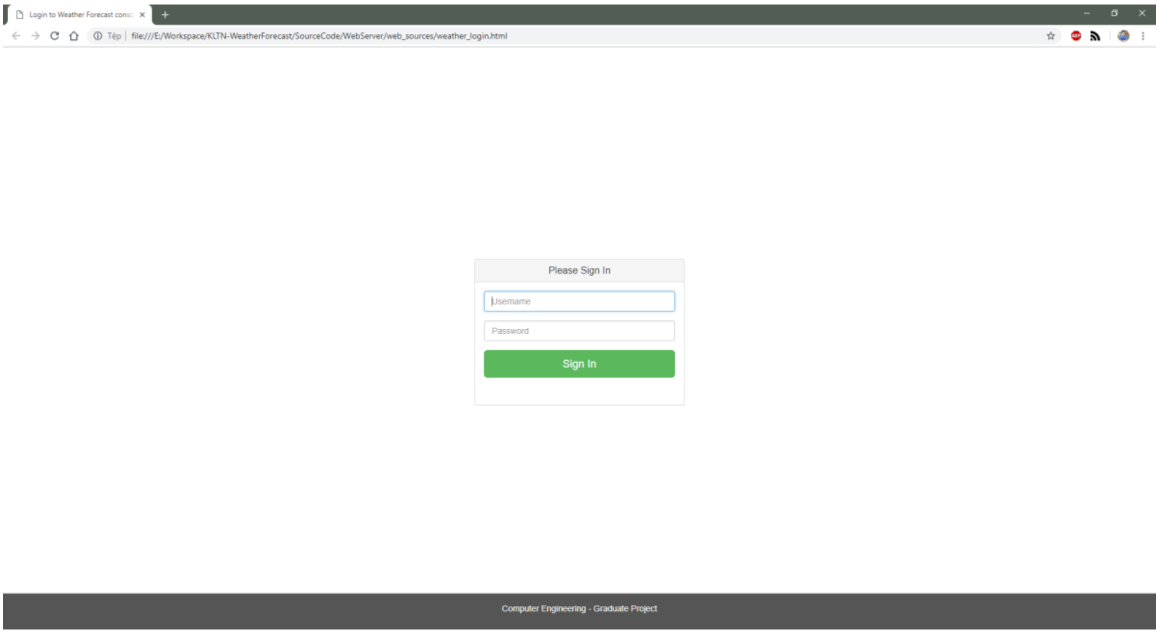


Figure 38: Login page

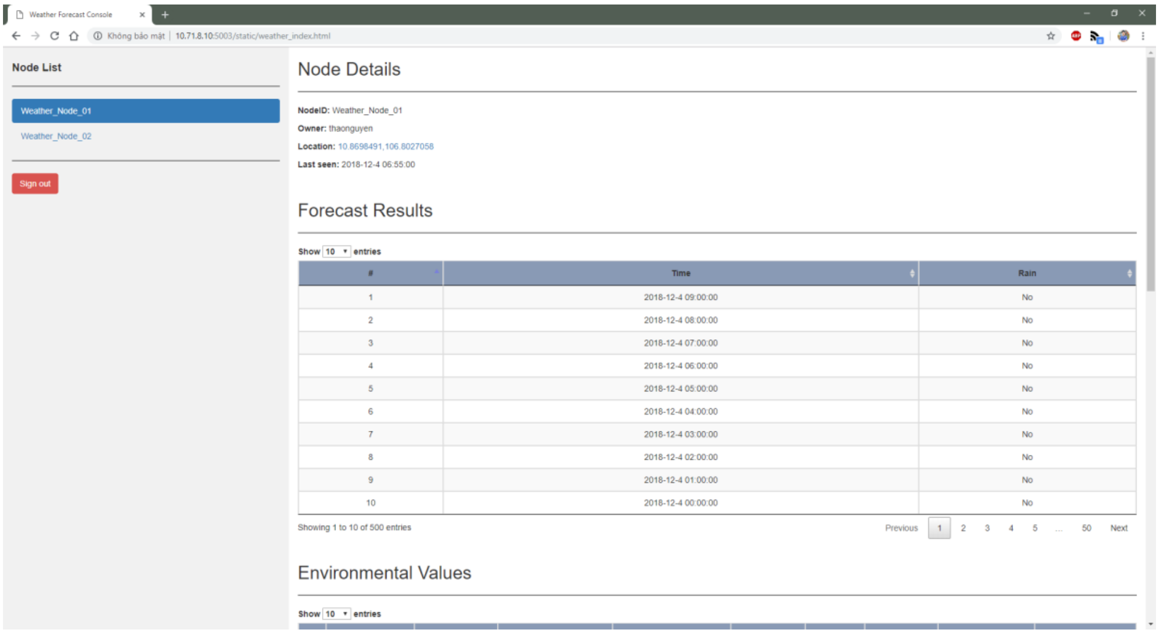


Figure 39: Data monitoring page

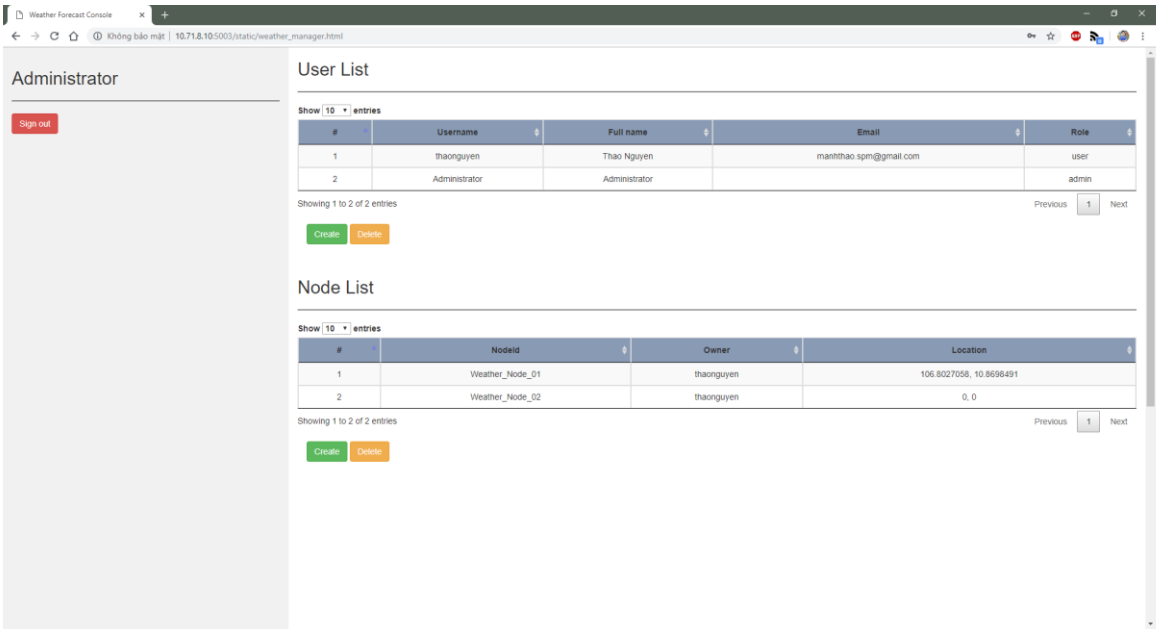


Figure 40: Account management page

3.4 Forecasting model design

To build the forecasting model, we divided the work into three main phases: sample data collection phase, data analysis and basic model building phase, integration of algorithms to improved outcomes phase.

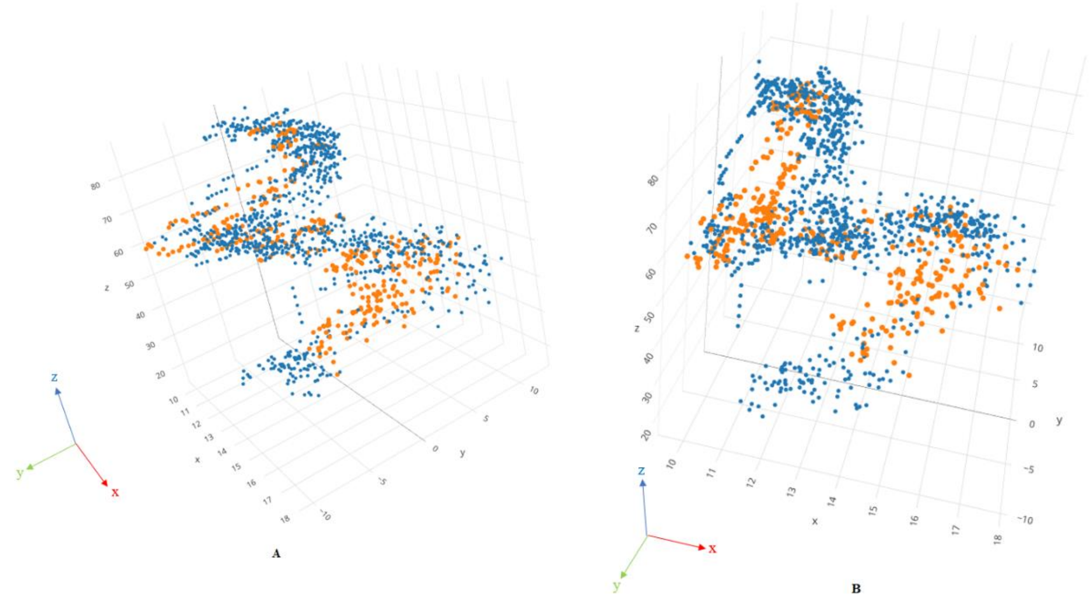
Phase 1: Sample data collection

At this stage, we used the first version of the Environmental Monitoring Node to collect data at the University of Information Technology. Data is collected and sent through Gateway and saved to Database in Server. The group conducted data collection from June to August 2018 and successfully collected about 21,000 records. Figure 41 shows the install location of Environmental Monitoring Node. In particular, each record was collected every 5 minutes and we expected to collect 25,900 records in 3 months. However, during the collection process, there are days when the device has a stability-related problem that does not collect data. Therefore, we only collected 21,000 records.

Figure 41: The environmental monitoring node installed on the roof of University of Information Technology HCMC

Phase 2: Data analysis and basic model building

After collecting sample data for 3 months, the team conducted data analysis and built the basic model from August 25^{th} to September 15^{th}, 2018.



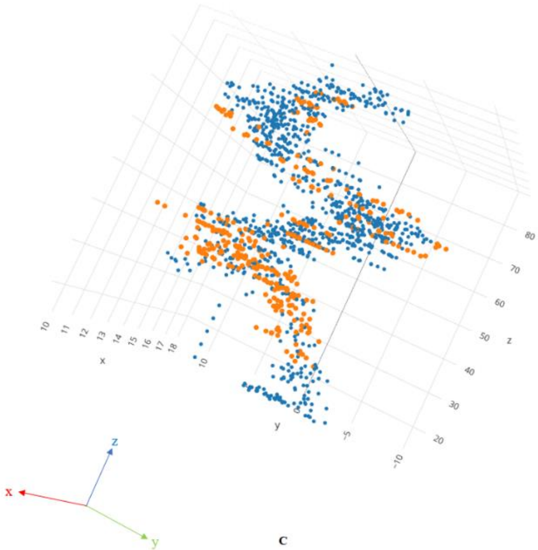


Figure 42: 3D plot of filtered sample data; A) front angle; B) Side angle; C) Above angle

Figure 42 shows different angles of our sample data filtered and plotted on 3D space. In the 3D plot, the orange points represent the rain and the blue points represent no-rain. We can see these points in space with a certain pattern and easy to classify.

From the sample data analysis, we began to build a basic forecasting model based on the ANN model. The reason why we choose the ANN model is that ANN suitable for the problem that we defined, it is strong enough to accurately predict the result, its accuracy is improved over time. In addition, the weather forecast is a long-term problem. Moreover, our problem is not big enough to use available large library models such as TensorFlow, etc.

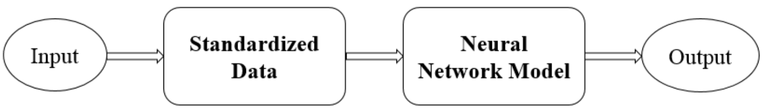


Figure 43: Forecasting model block diagram

Module 1: Standardized data: All weather parameters we collected are mapped to a corresponding value within the range [0, 1] :

- Wind speed, wind direction
- Time
- Lowest temperature in 1 hour
- Highest temperature in 1 hour
- Humidity is lowest in 1 hour
- Highest humidity in 1 hour
- Total pressure drops in 2 hours

(Based on expert opinion in the Agricultural Meteorological Curriculum Document of Agricultural University 1, in chapter XI, section 1.3b)

Module 2: Neural network: Inputs are from module 1, Output is the result of predicting whether there is rain or no rain. The transfer functions used are Sigmoid. [20].

Training data is a set of pairs (Xh, Yh), where X is the weather parameter of hour h in the day, Y is the corresponding expected result of the hour h + 2. Each pair of inputs set and desired output is called a training pattern. After being trained several times with such N training patterns, the neural network may have understood the rules of change of weather (in case N is large enough and training patterns represent the actual law of weather change in a season). The training process follows the reverse propagation algorithm shown above.

Result of this phase: We have successfully built the model and run the test to forecast over 20,000 records of sample data. However, the exact prediction results are not as accurate as expected.

Predict the causes:

- The output is not suitable for the model, we should divide the data set into 2 classes with rain (1), no rain (0).
- Input the lowest temperature and the highest temperature, the lowest humidity, and the highest humidity do not help the dataset split into 2

separate classes. Solution: average temperature in an hour, average humidity in an hour.

Phase 3: Integration of algorithms to improved outcomes

In this phase, we change the number of inputs from 8 to 6. Because we find that the minimum and maximum parameters of temperature and humidity do not help the model clearly show the classification in phase 2. These parameters are changed into 2 parameters: temperature reduction in an hour and temperature reduction in an hour. Also, at the output, we only want the result whether it is rain (1) or no-rain (0), we modify the transfer function at the output class from Sigmoid to Softmax [21]. Because Softmax function is strong in binary classification and it is widely used on other models.

During the implementation process, we discovered that the forecasting model encountered an overfitting error. Therefore, we apply more Dropout algorithms to the model to fix this Overfitting error.

There are 6 inputs:

- Wind speed
- Wind direction
- Time
- Total temperature reduction in 1 hour
- Total humidity reduction in 1 hour
- Total air pressure reduction in 2 hours

Output:

- Whether it’s rain

Model:

- Multi-layer perceptron with Sigmoid

- The transfer function is Sigmoid at hidden layers and Softmax Regression at the output layer
- Backpropagation and Dropout algorithms

From September 25th to December 2018, the forecasting model was improved, and the accuracy of the forecasting model was 90 percent. Specifically, these results will be presented in Chapter 4.

3.5 Anti-collision algorithm design

One of the problems with using LoRa technology is the packet collision. When two devices with the same configuration send packets at the same time, a packet error or even packet loss will occur.

There are many solutions from hardware to software algorithms such as Listen Before Talk, Channel Activity Detection, Sync Word, etc. However, these methods all have their own disadvantages and do not completely solve the problem. Especially with the RFM95W module used in the topic, the hardware is not strongly supported in handling packet collision.

Pure ALOHA and Slotted ALOHA both are the Random Access Protocols, that is implemented on the Medium Access Control (MAC) layer, a sublayer of Data Link Layer. The purpose of the ALOHA protocol is to determine that which competing station must get the next chance of accessing the multi-access channel at MAC layer. The main difference between Pure ALOHA and Slotted ALOHA is that the time in Pure Aloha is continuous whereas, the time in Slotted ALOHA is discrete.

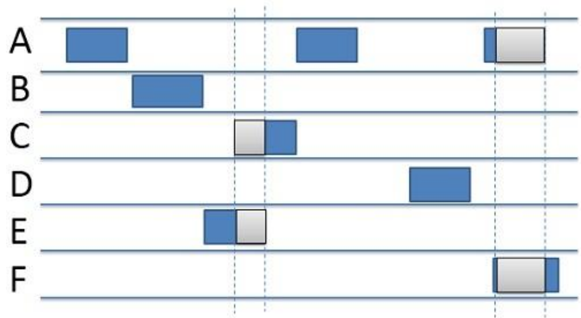


Figure 44: Pure ALOHA operations

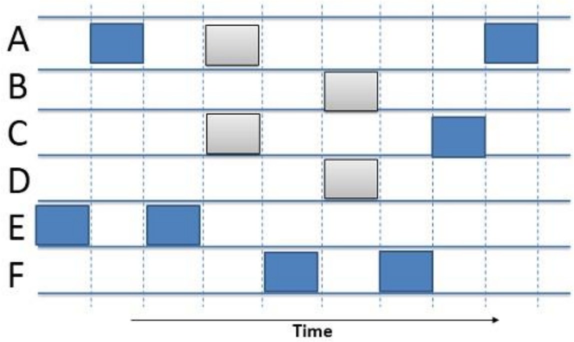


Figure 45: Slotted ALOHA operations

Anti-collision is an algorithm for resolving packet collisions including basic protocols such as sending back error packets or error messages, etc. Currently, LoRa has two common anti-collision algorithms which are Pure ALOHA and Slotted ALOHA. However, both aim to minimize latency when resending packets that fail and have the very low bandwidth. Pure ALOHA is 0.184 when G = 1/2 and Slotted ALOHA is 0.368 when G = 1 (G is the expected number of transmission and retransmission per unit time). So, both are inappropriate for weather monitoring systems, which send data every 5 minutes and latency is not important.

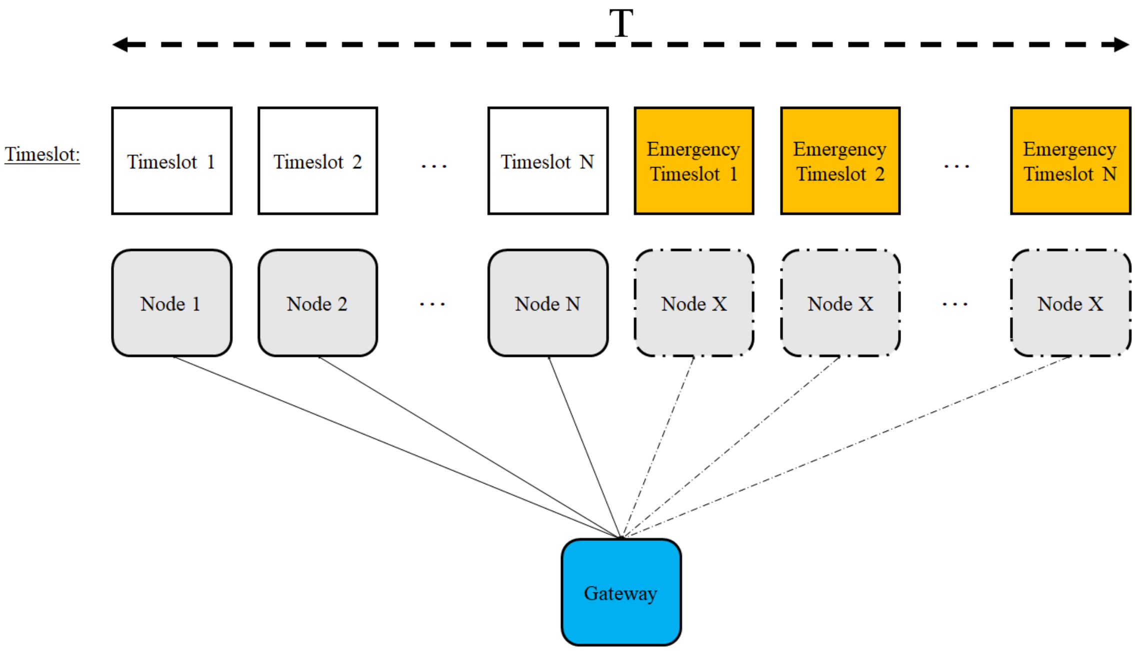


Figure 46: Anti-collision algorithm used in thesis

We have developed a new algorithm to solve this problem based on the current version of Slotted ALOHA algorithm. In this new algorithm, the transmissions are arranged in slots similar to Slotted ALOHA but have specific cycle time and emergency slots. The length of the cycle and each slot depends on the length of the packet or its time on air. The number of emergency slots usually accounts for about 20% of the total number of slots and depends on the LoRa packet configurations. Typically, slots will be used for sending the data packet. When a collision occurs, the nodes resend the packet with error information in the emergency slot. In addition, the emergency slot is also used to inform the server about system failure events. A disadvantage of this algorithm is the need for an RTC module on each device and a time synchronization slot.

Chapter 4. **RESULT AND SUMMARY**

4.1 Results

4.1.1 Forecasting accuracy

Forecast accuracy checking model

We have developed a model to test the accuracy of forecasts as follows:

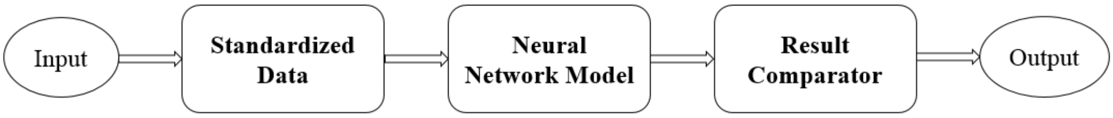


Figure 47: Forecast accuracy checking model

Input data goes through Standardized data blocks and is mapped to the range [0, 1]. After that, mapped data go to the neural network model to conduct the forecast. The result of the neural network model is the percentage of rain or no rain. The forecasting results will be compared with actual rainfall collected from the sensor. The comparisons are given in tables 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14.

In addition, we built 3 more models using different algorithms to compare with the model we implemented in this thesis. These models are Bayes [8], Decision Trees [8], SVM [22]. The forecasting results of these 3 models are similarly processed.

Achievements

The following tables show the comparison results in September, October and November 2018. The green cell indicates a prediction that match the actual result, whereas red indicates a wrong prediction. The last row of each table is the correct prediction rate for that day.

Table 3: Forecasting result of Bayes model in September 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
1	1	1	0	1	0	1	1	1	0	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	0	1	
2	1	0	0	1	0	1	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	1	0	0	1	
3	1	0	0	1	0	0	0	0	1	1	0	1	0	1	1	0	0	1	0	1	1	0	1	0	0	0	1	1	1	1	
4	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	1	0	0	0	1	0	1	1	
5	0	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	1	1	0	0	1	0	1	0	1	0	1	0	0	0	
6	0	1	0	0	1	1	0	0	0	0	1	0	1	0	0	1	1	1	1	0	0	1	1	0	1	0	1	0	0	0	
7	0	1	0	0	1	1	1	0	0	0	0	0	1	0	0	1	1	0	1	0	0	1	1	1	1	0	1	0	0	0	
8	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	0	1	0	1	1	0	1	0	1	1	0	0	1	0	0	
9	0	1	1	0	1	0	0	0	0	1	0	1	0	0	1	1	0	0	1	1	0	0	0	1	0	1	0	1	1	0	
10	1	0	1	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0	
11	1	0	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	
12	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	
13	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	0	0	1	
14	0	0	0	1	0	0	0	0	1	0	1	1	1	0	1	0	0	1	0	0	0	1	0	0	0	0	1	0	0	1	
15	0	0	1	0	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	0	0	1	0	0	1	1	1	0	0	1	
16	0	1	1	0	1	1	1	0	1	0	1	1	0	0	1	0	1	0	1	0	1	0	0	1	1	1	1	0	0	1	
17	0	1	1	0	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	1	1	0	0	1	1	1	1	0	0	0	
18	1	0	0	0	1	0	1	1	0	0	0	0	0	1	0	0	1	0	0	1	1	1	0	0	1	0	1	0	0	0	
19	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0	1	1	0	0	
20	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	
21	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	1	
22	0	1	1	1	0	0	1	0	1	1	1	0	1	0	0	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	
23	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	
	46	46	42	42	42	38	38	42	42	33	42	46	42	42	38	42	50	46	46	46	54	46	50	42	46	42	58	46	33	50	43.75

Table 4: Forecasting result of Decision tree model in September 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	0	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	
2	1	0	0	1	0	1	0	1	1	0	0	1	0	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	
3	1	0	0	1	1	0	0	1	1	1	0	1	0	1	1	0	0	1	0	1	1	0	1	0	1	1	1	1	1	1	
4	1	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	1	1	0	0	1	0	1	0	0	0	1	1	1	1	
5	0	1	0	1	0	0	0	0	1	0	1	0	1	1	0	0	1	1	1	0	1	0	1	0	1	0	1	0	0	1	
6	0	1	1	0	1	1	1	0	1	0	1	0	1	1	0	1	1	1	1	1	0	1	1	0	1	0	1	0	0	0	
7	0	1	1	0	1	1	1	0	0	0	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	1	0	0	0	
8	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	0	1	0	1	1	0	1	0	1	1	1	0	1	1	0	
9	1	1	1	0	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	1	0	
10	1	1	1	1	0	0	0	1	0	1	0	0	0	1	1	0	0	0	1	0	1	0	0	1	0	1	0	1	1	0	
11	1	0	1	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	0	
12	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	0	1	1	
13	1	0	0	0	1	0	0	0	1	1	1	0	1	0	0	1	0	1	0	0	0	1	1	0	0	0	1	0	1	1	
14	0	0	0	1	1	1	0	0	1	1	1	1	1	0	1	0	1	1	0	0	0	1	0	0	0	1	1	0	0	1	
15	0	1	1	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	0	0	1	0	1	1	1	1	0	0	1	
16	0	1	1	0	1	1	1	0	1	0	1	1	0	1	1	1	1	0	1	0	1	0	0	1	1	1	1	0	0	1	
17	0	1	1	0	1	1	0	1	0	0	0	1	0	1	0	0	1	1	0	1	1	0	0	1	1	1	1	0	0	0	
18	1	0	0	0	1	0	1	1	0	0	0	0	0	1	0	1	1	0	0	1	1	1	0	0	1	1	1	1	0	0	
19	1	0	0	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	1	0	1	1	0	1	0	1	1	0	0	
20	1	0	0	0	0	0	0	1	0	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	1	
21	0	0	0	1	0	0	0	0	0	0	1	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	0	1	0	1	
22	0	1	1	1	1	0	1	0	1	1	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	1	0	1	0	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	
	58	54	54	58	63	46	46	54	50	42	50	54	50	58	46	50	58	54	54	58	63	54	58	50	58	58	67	58	46	63	54.44

Table 5: Forecasting result of SVM model in September 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
2	1	1	1	1	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
3	1	1	0	1	1	1	1	0	1	1	0	1	0	1	1	0	1	1	0	1	1	0	1	0	1	1	1	1	1	1	
4	1	1	1	1	1	1	0	0	1	1	0	1	1	0	0	0	1	1	1	1	1	0	1	0	0	1	1	0	1	1	
5	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1	0	1	1	1	0	1	0	1	1	
6	0	1	1	1	1	1	1	0	1	0	1	0	1	0	0	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	
7	0	1	1	0	1	1	1	0	1	0	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	1	0	0	0	
8	0	1	1	0	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	1	0	1	0	1	1	0	1	1	1	0	
9	1	1	1	1	1	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	
10	1	1	1	1	0	1	0	1	0	1	1	0	0	1	1	1	1	0	1	0	1	1	0	1	1	1	0	1	1	1	
11	1	1	1	1	0	0	0	1	0	1	1	0	0	1	1	1	1	0	0	1	1	1	1	0	0	1	1	0	1	1	
12	1	0	0	1	0	0	0	1	1	1	1	0	0	0	0	1	0	1	0	1	1	1	1	1	0	0	1	0	1	0	
13	1	0	0	1	1	0	0	1	1	1	1	0	1	0	1	1	0	1	0	1	0	1	1	0	0	0	1	1	1	1	
14	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	0	1	
15	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	0	1	0	1	1	1	1	1	0	1	
16	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	
17	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	0	
18	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0	
19	1	0	0	1	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	
20	1	0	0	1	0	0	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	0	1	1	1	0	0	1	0	1	
21	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	75	75	75	83	79	79	75	75	71	67	71	79	75	71	71	71	79	83	75	71	79	71	83	71	71	71	83	71	79	75	75.14

Table 6: Forecasting result of Multi-layer perceptron model in September 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
2	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
3	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	
4	1	1	1	1	1	1	0	0	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1	0	1	1	
5	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	
6	0	1	1	1	1	1	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
7	0	1	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	
8	0	1	1	0	1	1	1	0	0	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	0	1	1	1	0	
9	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
10	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	
11	1	1	1	1	0	0	0	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	
12	1	0	0	1	1	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
13	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
14	1	0	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
18	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
19	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
20	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	
21	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	87.5	83.3	83.3	91.7	91.7	91.7	79.2	83.3	91.7	75	87.5	75	100	95.8	79.2	100	100	83.3	95.8	91.7	91.7	91.7	100	95.8	83.3	87.5	95.8	87.5	91.7	79.2	89.03

Table 7: Forecasting result of Bayes model in October 2018

	Day																															
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	
2	1	0	0	1	1	1	1	1	0	1	0	1	1	0	1	0	1	0	0	1	0	1	1	0	1	1	1	0	1	1	0	
3	1	0	0	1	0	0	1	0	0	0	0	1	1	0	1	0	1	1	1	1	0	0	1	0	0	1	0	0	1	0	1	
4	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	0	0	1	0	0	1	0	0	1	0	1	
5	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	0	1	1	0	1	
6	0	0	1	1	0	0	0	0	0	0	1	1	0	1	0	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	
7	0	1	0	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	
8	0	1	0	0	1	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	1	0	1	1	0	1	0	1	0	0	
9	1	0	0	0	1	0	0	1	1	0	0	1	0	0	0	1	0	0	0	0	0	1	0	1	0	0	1	0	1	0	0	
10	1	0	0	0	0	1	0	1	1	1	0	0	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	0	0	0	0	
11	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	1	1	0	0	1	0	1	0	0	1	0	0	0	0	0	
12	0	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	0	1	1	0	0	0	0	1	0	1	0	0	1	
13	0	0	0	0	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	1	0	1	
14	0	0	0	1	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	0	1	
15	0	0	0	1	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0	1	1	0	0	1	
16	0	1	0	1	0	0	1	0	1	0	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	0	
17	1	1	0	0	1	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	0	
18	0	1	1	0	1	0	1	0	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	1	0	1	0	0	0	1	0	
19	0	1	1	0	1	0	1	0	0	1	1	0	0	0	1	0	0	0	1	0	1	0	1	1	0	1	0	0	0	1	0	
20	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	1	1	
21	0	0	1	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	
22	0	1	1	0	0	1	1	0	0	0	0	1	1	0	1	0	1	0	1	0	1	0	0	0	0	0	0	1	1	0	0	
23	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1	0	1	0	1	1	0	0	
	29	42	46	46	33	46	46	29	33	38	38	54	38	42	50	33	42	42	42	38	38	29	46	42	42	50	38	46	50	38	50	40.99

Table 8: Forecasting result of Decision tree model in October 2018

	Day																															
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	
2	1	0	0	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	0	1	0	1	1	0	1	1	1	0	1	1	1	
3	1	0	0	1	0	1	1	1	0	1	0	1	1	0	1	0	1	1	1	1	0	0	1	0	1	1	0	0	1	0	1	
4	0	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	1	1	1	0	0	0	1	0	0	1	0	0	1	0	1	
5	0	0	1	1	0	0	0	0	0	0	1	1	0	0	0	0	0	1	1	1	0	0	0	0	1	1	0	1	0	0	1	
6	0	0	1	1	0	0	0	0	1	0	1	1	0	1	0	1	0	0	1	0	0	0	0	1	1	0	0	1	0	0	1	
7	0	1	1	1	1	0	0	0	1	0	1	1	0	1	0	1	0	0	1	0	0	1	0	1	1	0	1	1	0	1	0	
8	0	1	1	0	1	0	0	0	1	0	0	1	0	1	0	1	0	0	0	0	0	1	0	1	1	0	1	0	1	1	0	
9	1	1	0	0	1	1	0	1	1	0	0	1	0	1	1	1	1	0	0	0	0	1	0	1	0	1	1	0	1	1	0	
10	1	0	0	0	0	1	0	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	0	0	0	0	
11	1	0	0	0	0	1	0	1	0	1	0	0	1	0	0	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	1	
12	0	0	0	0	0	1	0	1	0	1	0	0	1	0	0	0	1	1	0	1	1	0	1	0	1	1	0	1	0	0	1	
13	0	0	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	1	
14	0	0	0	1	0	1	1	0	1	0	1	0	1	1	0	0	0	1	0	1	0	0	0	0	1	0	0	1	1	0	1	
15	0	1	0	1	0	0	1	0	1	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	1	0	1	1	0	1	1	
16	1	1	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	1	1	0	1	0	
17	1	1	1	1	1	0	1	0	0	1	0	0	0	1	1	1	1	1	0	0	0	1	1	1	0	0	1	0	0	1	0	
18	1	1	1	0	1	0	1	0	0	1	1	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	0	0	0	1	0	
19	0	1	1	0	1	0	1	0	0	1	1	1	0	0	1	1	0	0	1	0	1	1	1	1	0	1	0	0	0	1	0	
20	0	0	1	0	1	0	1	1	0	0	1	1	0	0	1	1	0	0	1	0	1	1	0	1	0	1	0	1	1	1	1	
21	0	0	1	0	0	1	1	1	0	0	0	1	0	0	1	0	0	0	1	0	1	0	0	1	0	0	1	0	1	1	1	
22	0	0	1	1	0	1	1	0	1	0	0	1	1	0	1	0	1	0	1	1	1	0	1	1	1	0	1	1	1	0	0	
23	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	
	46	46	58	63	42	54	58	50	46	50	46	63	42	54	58	46	54	54	54	46	42	50	54	58	58	58	54	50	54	54	63	52.42

Table 9: Forecasting result of SVM model in October 2018

Hour \ Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	
2	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	
3	1	0	1	1	0	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	1	0	1	1	0	1	1	0	1	
4	1	0	1	1	1	0	0	1	0	1	1	0	0	0	1	0	1	1	1	0	1	0	1	0	0	1	0	1	1	1	1	
5	0	1	1	1	0	1	0	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	
6	0	1	1	1	0	1	0	0	1	0	1	1	0	1	0	1	0	0	1	1	0	0	0	1	1	0	1	1	0	1	1	
7	1	1	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	0	1	1	0	1	0	1	1	0	1	1	0	1	0	
8	1	1	1	0	1	0	1	0	1	1	0	1	1	1	0	1	0	0	0	1	0	1	1	1	1	0	1	0	1	1	0	
9	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	
10	1	0	0	0	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	0	
11	1	0	0	0	0	1	0	1	0	1	0	0	1	0	1	1	1	1	0	0	1	1	1	0	0	1	0	1	0	0	1	
12	0	1	1	1	1	1	0	1	0	1	0	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	0	1	0	1	1	
13	0	1	1	1	1	1	0	1	0	1	1	0	1	0	1	0	0	1	1	1	0	1	0	0	1	1	1	1	1	0	1	
14	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	0	1	1	1	1	0	0	1	1	1	0	0	1	1	0	1	
15	0	1	0	1	0	1	1	0	1	0	1	1	1	1	0	0	0	1	0	1	0	0	0	1	1	0	1	1	1	1	1	
16	1	1	0	1	0	1	1	0	1	0	0	1	0	1	0	1	0	1	0	0	1	1	0	1	1	0	1	1	1	1	1	
17	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	0	
18	1	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	0	1	0	
19	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	0	1	0	1	0	
20	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
21	1	0	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	1	0	1	1	1	
22	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	
	71	71	75	83	67	75	75	71	71	71	67	79	75	67	75	71	71	71	71	71	71	71	75	67	75	71	71	75	71	75	75	72.31

Table 10: Forecasting result of Multi-layer perceptron model in October 2018

Hour \ Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	
2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	
4	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	
5	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	
6	0	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	
8	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	
9	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	
10	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	
11	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	
12	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	
13	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	
14	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	91.7	100	91.7	83.3	100	100	91.7	95.8	100	100	100	87.5	100	91.7	87.5	100	91.7	83.3	83.3	95.8	95.8	91.7	91.7	83.3	95.8	87.5	95.8	95.8	100	95.8	95.8	93.68

Table 11: Forecasting result of Bayes model in November 2018

Hour \ Day	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	0	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	
1	1	0	1	1	1	1	0	1	0	1	1	0	1	0	1	1	0	1	1	0	1	0	1	1	1	0	1	1	1	1	
2	0	0	1	0	1	0	1	1	1	0	1	0	1	0	0	1	1	0	1	0	1	1	0	1	1	1	1	1	1	0	
3	0	0	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	0	1	1	1	1	1	0	1	1	0	1	0	1	
4	0	0	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	1	
5	0	1	0	0	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	0	0	1	0	1	1	1	1	1	1	
6	0	1	0	1	0	0	0	0	0	0	1	1	1	1	0	1	0	1	0	0	1	1	1	0	1	0	1	0	0	1	
7	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	1	1	1	0	0	0	1	0	1	0	0	1	0	0	1	
8	1	1	0	0	1	0	0	0	0	1	0	0	0	1	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	
9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	0	1	
10	1	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	
11	1	0	0	0	1	1	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
12	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
13	0	1	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	
14	0	1	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	
15	0	1	0	0	0	1	1	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	
16	1	1	0	1	0	1	0	0	1	1	0	1	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	1	0	
17	1	1	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	1	1	1	0	0	0	1	1	0	0	0	1	0	
18	1	1	0	1	1	0	0	1	1	0	0	1	0	1	0	1	0	1	1	1	0	0	0	1	1	1	0	0	1	1	
19	1	1	0	1	1	0	0	1	1	0	0	0	0	1	0	1	1	1	1	1	0	0	1	1	0	1	0	1	1	1	
20	1	0	1	1	1	0	0	1	1	0	0	0	0	1	0	1	1	0	1	1	0	1	1	0	0	1	1	0	0	1	
21	1	1	0	0	1	0	1	1	0	0	1	1	1	1	0	1	1	1	0	0	1	1	1	0	0	0	1	1	0	1	
22	0	1	1	0	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	0	1	0	0	1	1	1	0	0	
23	0	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	1	0	1	0	0	0	1	1	0	1	
	50	63	33	46	58	33	42	50	33	46	58	46	42	42	33	54	50	46	42	46	46	33	58	42	54	38	42	42	46	54	45.56

Table 12: Forecasting result of Decision tree model in November 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	
1	1	0	1	1	1	1	0	1	1	1	1	0	1	0	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	
2	1	0	1	0	1	0	1	1	1	0	1	0	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	
3	0	0	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	1	1	0	1	1	0	1	0	1	
4	0	0	1	1	1	0	1	1	1	0	1	1	1	0	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	
5	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	1	0	1	1	1	0	0	1	0	1	0	1	1	0	1	
6	0	1	0	1	0	0	0	0	0	1	0	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	0	0	1	
7	1	1	0	0	0	0	0	0	0	1	0	1	0	1	0	1	1	1	0	0	1	1	0	1	1	0	1	0	0	1	
8	1	1	0	0	1	0	0	0	0	1	0	1	0	0	1	0	1	0	0	0	1	0	0	1	0	0	0	0	1	0	
9	1	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	1	0	0	0	0	1	1	
10	1	0	1	0	1	1	0	1	0	0	1	0	0	0	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	
11	1	0	1	0	1	1	0	0	0	0	1	0	0	0	1	0	1	1	1	0	0	0	0	1	0	1	0	0	0	0	
12	0	0	1	0	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0	0	0	1	0	0	1	1	0	0	0	0	
13	0	0	0	1	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	0	
14	0	1	0	1	0	0	1	0	0	1	1	1	0	1	0	0	0	0	0	1	1	1	0	0	1	0	1	1	1	1	
15	0	1	0	0	0	1	1	0	1	1	0	1	0	1	0	0	0	0	0	1	1	0	0	0	1	0	1	0	1	1	
16	1	1	0	0	0	1	1	1	1	1	0	1	0	0	0	1	0	1	0	1	1	0	1	1	1	0	0	0	1	1	
17	1	1	0	1	1	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	0	0	0	1	1	1	0	0	1	0	
18	1	1	1	1	1	0	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0	1	1	
19	1	1	1	1	1	0	0	1	1	0	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	
20	1	1	1	1	1	0	1	1	0	0	1	1	0	1	0	1	1	0	1	1	0	1	1	1	0	1	1	0	0	1	
21	1	1	0	0	1	0	1	1	0	0	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	0	1	1	0	1	
22	0	1	1	0	1	0	1	1	0	1	1	0	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0	0	
23	0	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	1	1	1	0	0	1	1	1	0	1	
	58	63	58	50	67	42	50	67	38	54	63	58	42	50	46	50	63	67	54	63	63	54	63	54	67	54	54	50	50	67	55.83

Table 13: Forecasting result of SVM model in November 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	
1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	
2	0	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	
3	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	1	
4	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	
5	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	0	1	0	1	1	1	1	
6	0	1	0	1	0	1	0	1	1	1	0	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	
7	1	1	1	1	0	0	0	1	1	1	1	1	0	1	0	1	1	1	0	1	1	1	0	1	0	0	1	1	0	1	
8	1	1	0	0	1	0	1	0	1	1	0	0	0	0	1	1	1	0	0	1	1	1	0	1	0	0	0	1	1	1	
9	1	1	1	0	1	0	1	1	1	0	0	0	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	
10	1	0	1	1	1	1	0	1	1	0	1	1	0	1	1	1	1	0	1	1	0	1	0	1	0	1	0	0	0	1	
11	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	1	0	1	1	1	0	0	
12	0	1	0	0	1	1	0	0	0	0	1	0	1	0	1	0	0	1	1	0	1	0	1	0	1	0	1	0	1	0	
13	0	1	0	1	0	0	1	0	0	1	1	0	1	0	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	0	
14	0	1	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	
15	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	1	1	1	
16	1	1	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	
17	1	1	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
18	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
20	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	
21	1	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	
22	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	
23	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
	75	88	75	71	75	75	71	88	79	79	79	75	79	71	75	88	75	75	75	88	83	79	79	75	75	75	79	88	75	79	78.06

Table 14: Forecasting result of Multi-layer perceptron model in November 2018

	Day																														
Hour	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
2	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
5	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
6	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	
7	1	1	0	1	0	1	0	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
8	1	1	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	
12	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
14	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	
15	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
22	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
23	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	
	100	100	87.5	100	91.7	87.5	83.3	100	100	100	95.8	95.8	91.7	100	100	100	91.7	100	100	100	87.5	100	100	83.3	83.3	100	100	100	100	91.7	95.69

Through the results above, we can see that the Multi-layer Perceptron model with Back-propagation algorithms produces more accurate results than other models.

Specifically, the average accuracy that the group conducted is 92.8%, 75.17%, 54.23% and 43.43% corresponding to Multi-layer Perceptron, SVM, Decision tree and Bayes models. In particular, in the region where we installed device still has a lot of rain and special weather in September such as storms and tropical depressions should have lower results than other months.

From the above results, the longer ANN model works, the more accurate it will be. In general, false predictions focus on periods of rainy weather. We have learned that the model has not been completed, the predictions at the accuracy record rain are not high. The reason is that the sample data collected is mainly concentrated in the last months of the rainy season in southern Vietnam, it does not rain or rain very little, leading the model to have a deflection toward the direction without much rain.

4.1.2 Environmental monitoring node power consumption

We recorded the power consumption at stages and give the following comparison table. The 2016 version is the original version of NGUYEN TRAN TIEN DAT and HO QUI DAY. The Atmega version is the first version of us, removing some unnecessary modules from the old version. The STM8 version is the final version, using STM8 microcontrollers and adding some other battery-related features such as measuring the battery's remaining capacity or allowing battery charging.

We measured the power consumption of the versions or DUT (Device under test) according to the diagram in Figure 48 and gave the results in Table 15.

Table 15: Power consumption on environmental monitoring node

Version	Voltage	Current	Power	Note
2016	5V	23.4mA	117mW	
Atmega	5V	17.3mA	86.5mW	
Atmega	3.3V	9.3mA	30.69mW	Sleep modules
STM8	9V	11.5mA	103.5mW	
STM8	9V	250mA	2250mW	Charging

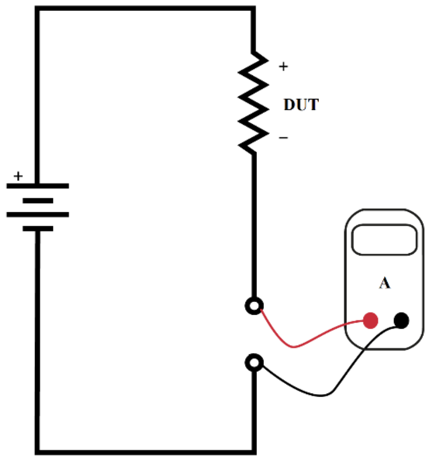


Figure 48: Diagram of power consumption measuring

Table 15 shows the results of power consumption measurement on different versions of environmental monitoring node. Versions operate at different voltages. To be comparable, we use formula (1) to calculate the power of each

$$P = U * I \qquad (1)$$

With P is Power, U is the voltage that the circuit operates, I is the current.

We tested all versions with 2 Panasonic 18650 5600mAh battery meets the requirement of stable and continuous operation. With the final version using STM8, the team tested with the battery fully charged and used it completely without charging, the system operated for more than 1 week from December 19 to December 27, 2018. Battery wear is also limited due to the charging allowing feature, so we do not need to replace the battery regularly.

Although the actual measurement results are many times higher than the theory, thanks to solar cells the system still satisfies the operation in stormy conditions in many days. Compared to the original version in 2016, the group has improved its energy usage by nearly 74% with the Atmega version of the modules sleeping.

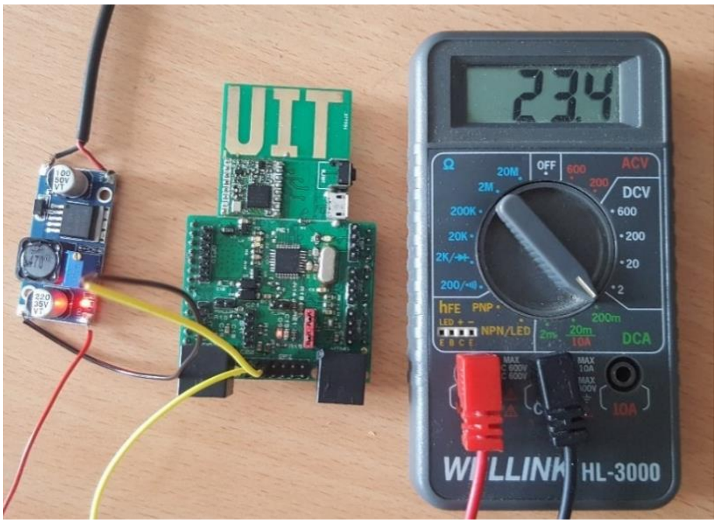


Figure 49: Power consumption on the version in 2016

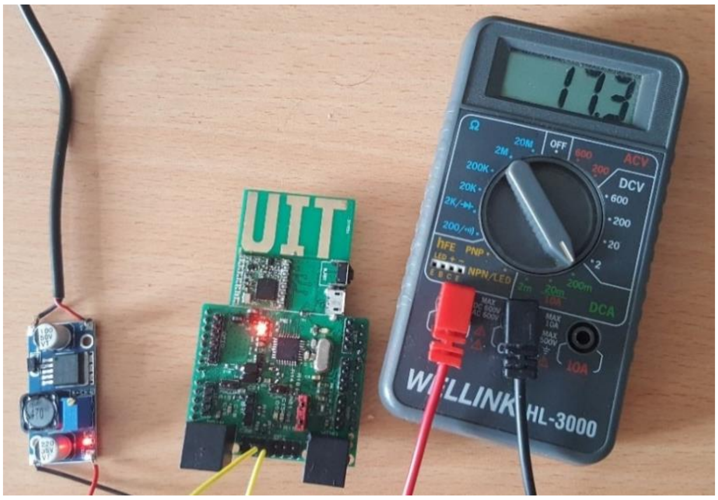


Figure 50: Power consumption on the Atmega version without Sleeping strategy

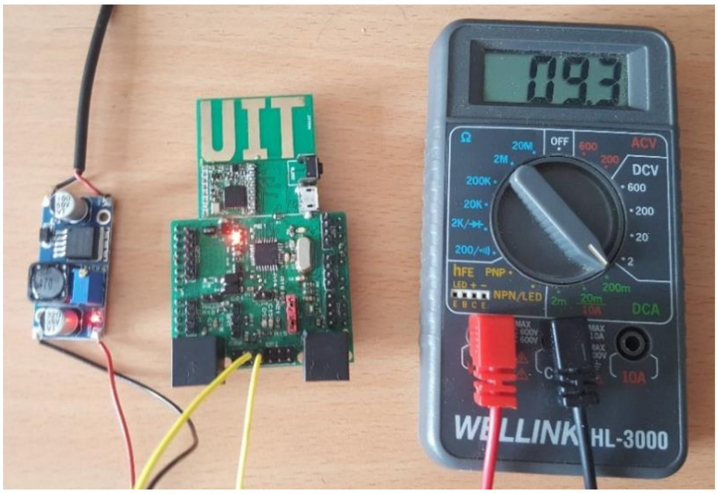


Figure 51: Power consumption on the Atmega version with Sleeping strategy

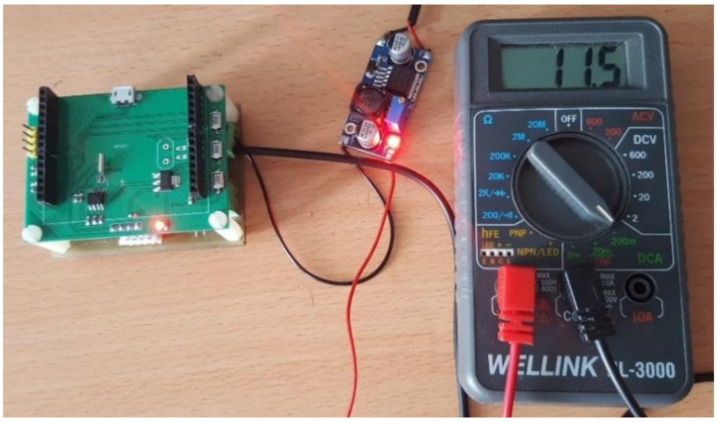


Figure 52: Power consumption on the final version using STM8

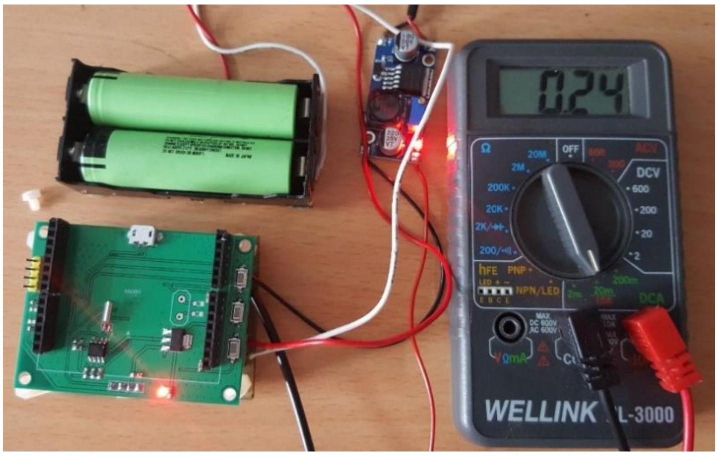


Figure 53: Power consumption on the final version in charging mode

4.1.3 LoRa anti-collision algorithm

We have conducted simulations with different settings to determine the appropriate parameters for each LoRa configuration. Specifically, we changed the Spreading Factor from SF7 to SF12 and tested on different Timeslot lengths. Other parameters of LoRa like Bandwidth, coding rate, and payload length are not changed through tests. The configuration of these parameters are:

- Bandwidth: 125k
- Coding rate: 4/5
- Payload length: 150 bytes

In each test, 20000 packets were sent. The following are the results:

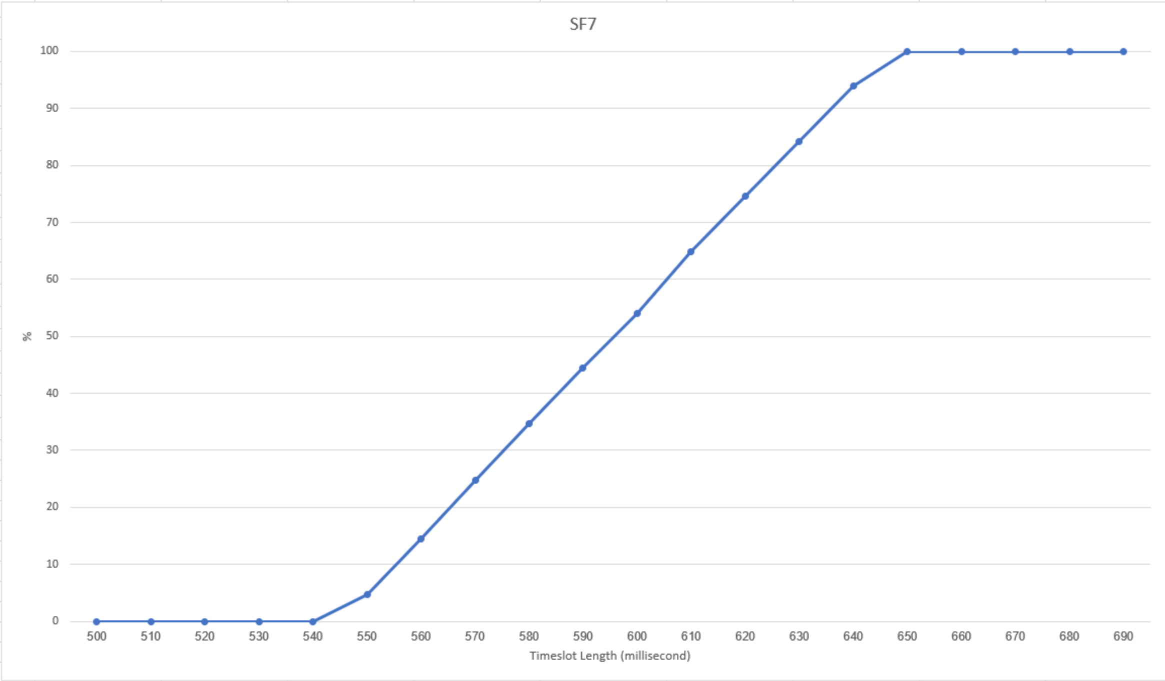


Figure 54: Successfully packet rate in SF7

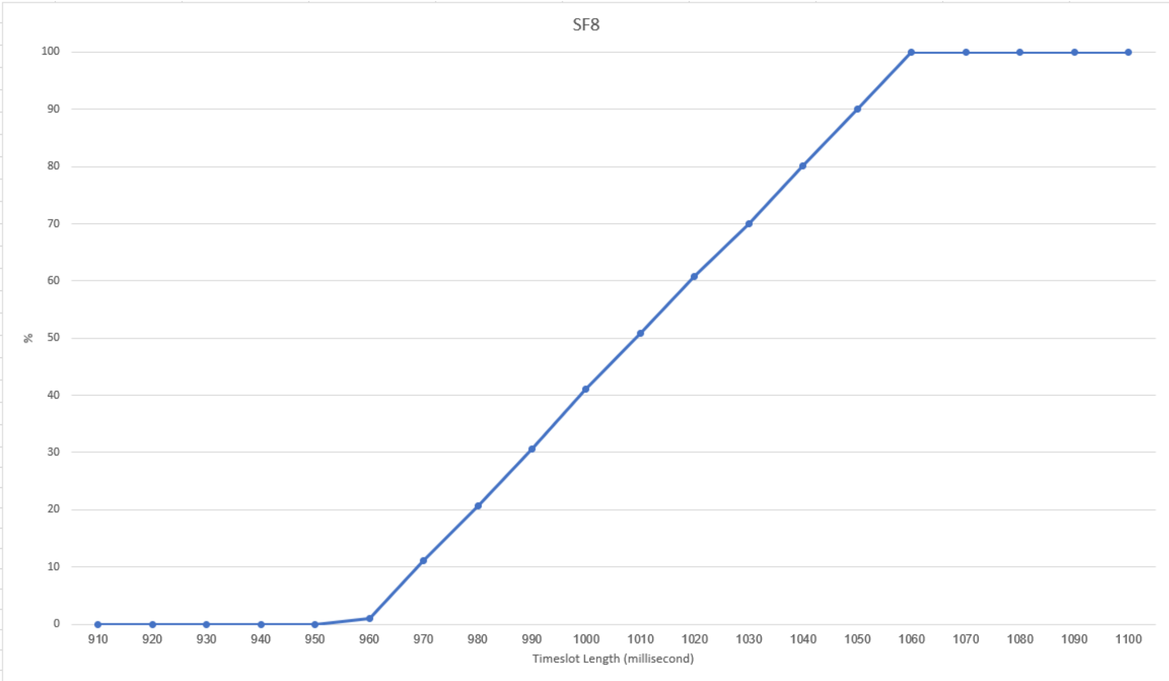


Figure 55: Successfully packet rate in SF8

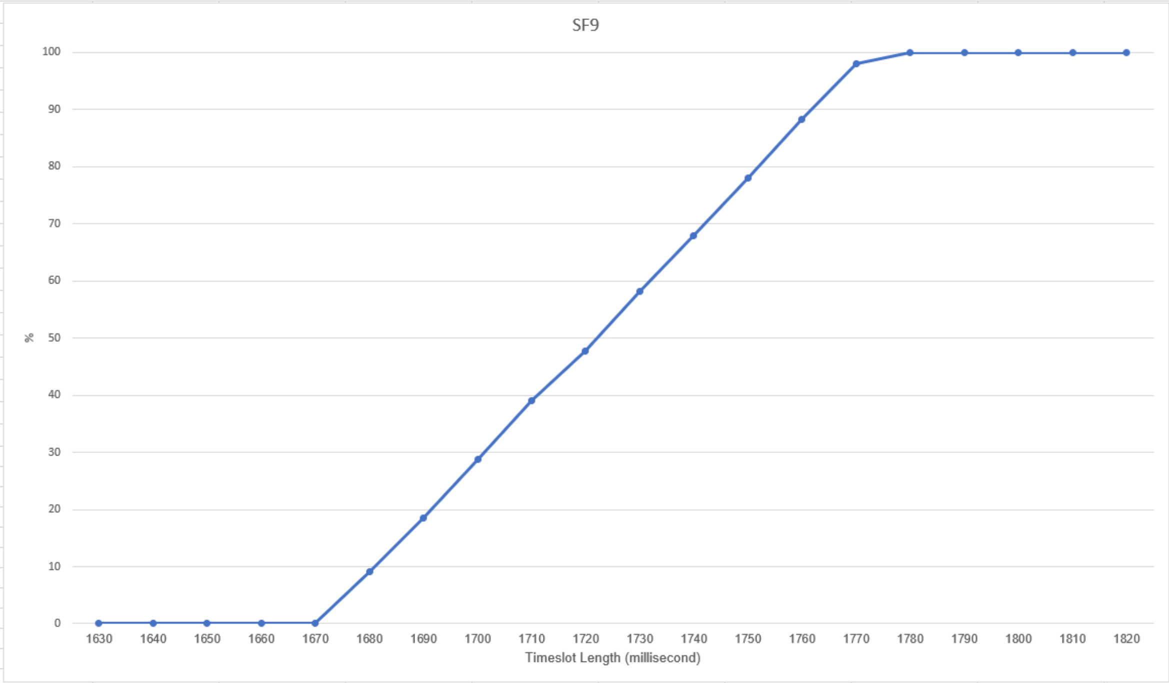


Figure 56: Successfully packet rate in SF9

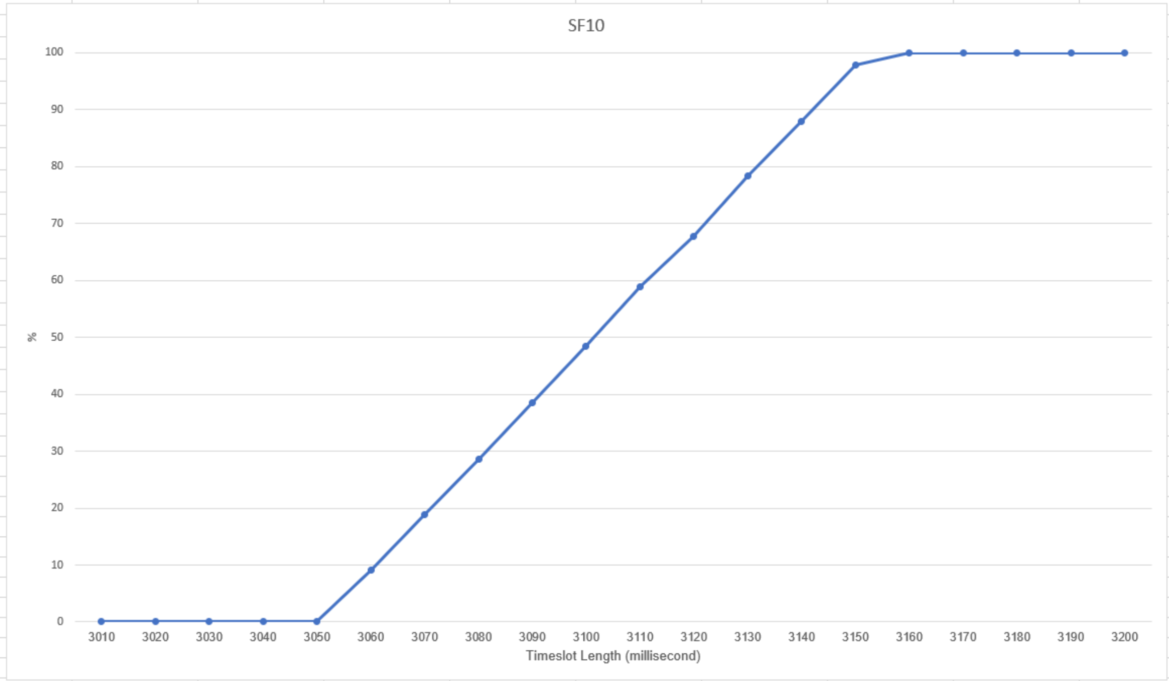


Figure 57: Successfully packet rate in SF10

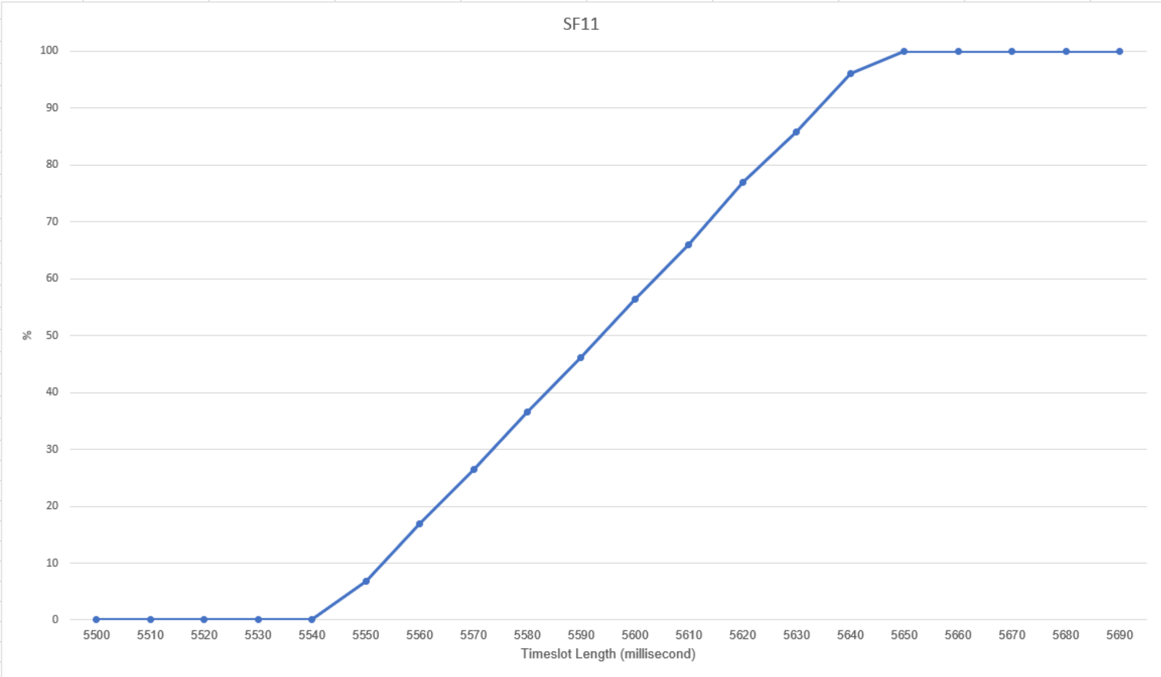


Figure 58: Successfully packet rate in SF11

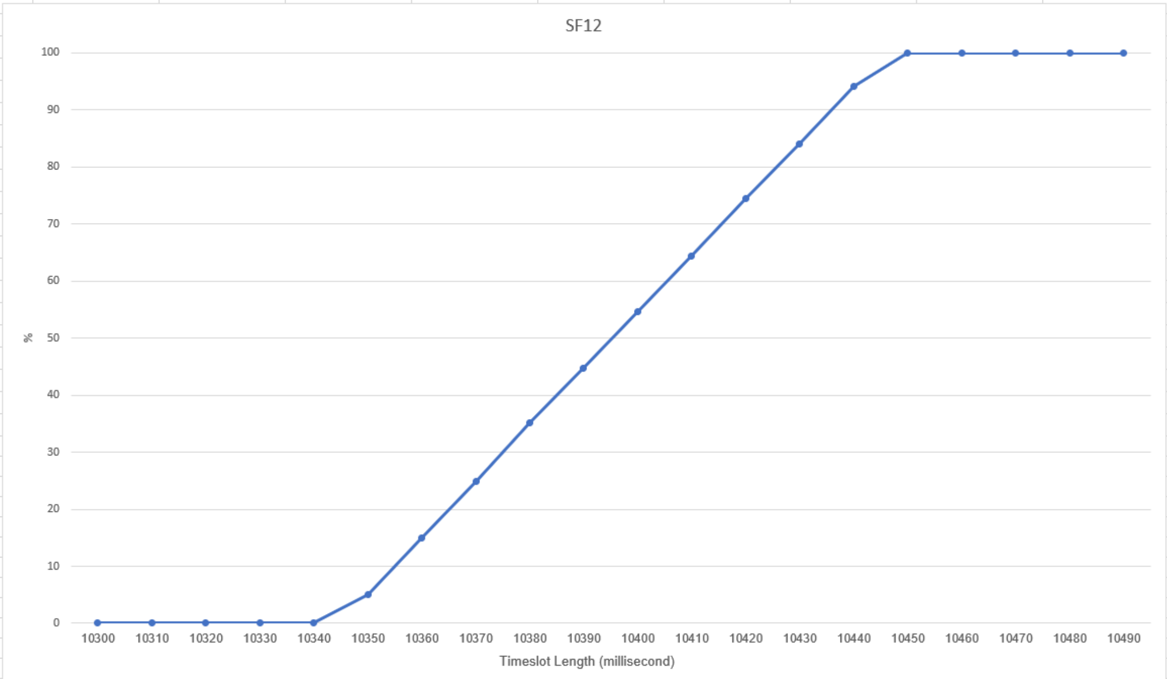


Figure 59: Successfully packet rate in SF12

Through the above results, we can see that the algorithm can respond well to all LoRa Spreading Factor by changing Timeslot length. Although increasing Spreading Factor in LoRa can increase the transmission distance and reduce the possibility of falling packages due to low receiver RSSI, the increasing in Speaking

Factor leads to Timeslot length increasing and the device will be slow update. Therefore, we suggested optimizing Spreading Factor with the transmission distance before applying Anti-collision algorithm.

Table 16: Timeslot length in different LoRa Spreading Factors

Spreading Factor	SF7	SF8	SF9	SF10	SF11	SF12
Timeslot Length (ms)	650	1060	1780	3160	5650	10450
Period (s)	39	63.6	106.8	189.6	339	627

Table 16 represents the length of a timeslot (in milliseconds) and the period of sending a data packet. Currently, our system is using SF11 (node sends data every 339 seconds or about 5 minutes) to achieve the necessary communication distance.

4.2 Summary

In general, the thesis has achieved the initial criteria such as:

- Successful development of rain forecasting system in small areas, independent and stable operation even in stormy weather
- Implement machine learning technology to improve the accuracy of prediction according to system usage time
- Using LoRa data transmission technology to apply to signal transmission, increase system installation distance
- The result of the prediction accuracy is also very positive when reaching nearly 90% of the exact predicted rain rate, much higher than the initially expected result.
- Optimizing energy consumption on monitoring equipment, increasing the ability to operate stably
- Regarding the Anti-collision algorithm, the team successfully built and applied to the topic, solving the problem of packet error or loss due to packet collision. Provide simulation results at different settings of LoRa.

However, the topic still has some problems as follows:

- The accuracy of prediction is high, but the predicted time is just 2 hours due to lack of input from the environment. To be able to predict over 2 hours, pictures of cloud from the satellite are required.
- Due to the time limit of 6 months, the sample data collection time is still short, only predicting some basic weather phenomena. More complex weather such as storms and tropical depressions need more sample data for training.
  - Giving the accuracy parameters gradually increased over the months but may be due to the end of the rainy season in South Vietnam so it has not been determined.

Chapter 5. **FUTURE WORK**

- Design node more compact, aesthetics and save more energy
- Based on weather forecasting theory and experimental study, the dramatic decrease of pressure leads to rain in 4 hours after. However, to forecast rainfall in more than 4 hours in advanced, picture of the cloud is required.
- Increase the number of environmental monitoring node to ensure that the data collected is exactly and enough
- Develop the system into a network so that it can monitoring weather of a bigger region
- Add OTA firmware update for Node and Gateway to improved usability and reduced costs of manual firmware upgrading
- Add Ethernet connection to Gateway, ensure usability for users
- Develop Self Hardware Failure Diagnostic system, improved hardware stability, especially environmental monitoring nodes when operating continuously outdoors and the system will notify when hardware has a problem

REFERENCES

[1] Saktaya Suksri and Warangkhana Kimpan, "Neural Network training model for weather forecasting using Fireworks Algorithm", 2016 International Computer Science and Engineering Conference (ICSEC). [Online]. Available: https://ieeexplore.ieee.org/document/7859952/. [Accessed Jul. 14, 2018].

[2] Andrew Culclasure, "Using Neural Networks to Provide Local Weather Forecasts", Electronic Theses & Dissertations Georgia Southern University Spring 2013. [Online]. Available: https://digitalcommons.georgiasouthern.edu/etd/32/. [Accessed Jul. 14, 2018].

[3] Kumar Abhishek, "Weather Forecasting Model using Artificial Neural Network", Procedia Technology Volume 4, 2012, Pages 311-318. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221201731200326X. [Accessed Jul. 14, 2018].

[4] Nguyễn Trần Tiến Đạt, Hồ Quí Đầy, "Thiết Bị Quan Trắc, Dự Báo Thời Tiết Phạm Vi Nhỏ Sử Dụng Công Nghệ Lora", Đại học Công nghệ Thông tin, 2016.

[5] Ram Niwas and Surender Singh, A Text Book on Agricultural Meteorology. Hisar, India: CCS Haryana Agricultural University, 2006.

[6] Đoàn Văn Điếm, Giáo trình Khí tượng Nông nghiệp. Hà Nội: Trường đại học Nông nghiệp 1, 2008.

[7] T. Mahmud, M. Hasan, A. Chakraborty, and A. K. Roy-Chowdhury, "A poisson process model for activity forecasting," 2016 IEEE International Conference on Image Processing (ICIP), 2016.

[8] Peter Norvig and Stuart J. Russell, Artificial Intelligence: A Modern Approach. Prentice Hall, 1994.

[9] Semtech, "What is LoRa®?", What is LoRa? | Semtech LoRa Technology | Semtech. [Online]. Available: https://www.semtech.com/lora/what-is-lora. [Accessed: 07-Jul-2018].

[10] The Things Network, "LoRaWAN", The Things Network, 17-Dec-2018. [Online]. Available: https://www.thethingsnetwork.org/docs/lorawan/. [Accessed: 07-Jul-2018].

[11] Lora Alliance, "About LoRaWAN™", LoRaWAN™. [Online]. Available: https://lora-alliance.org/about-lorawan. [Accessed: 07-Jul-2018].

[12] OASIS, "MQTT Version 3.1.1", MQTT Version 3.1.1. [Online]. Available: http://docs.oasis-open.org/mqtt/mqtt/v3.1.1/os/mqtt-v3.1.1-os.html. [Accessed: 15-Jul-2018].

[13] David A. Wheeler, Secure Programming for Linux and Unix HOWTO, v3.010 ed. 2003. GNU Free Documentation License

[14] MongoDB Inc, "What Is MongoDB?", MongoDB. [Online]. Available: https://www.mongodb.com/what-is-mongodb. [Accessed: 20-Jul-2018].

[15] "Weather Station with Anemometer/Wind vane/Rain bucket SKU:SEN0186 - DFRobot Electronic Product Wiki and Tutorial: Arduino and Robot Wiki-DFRobot.com", Dfrobot.com, 2019. [Online]. Available: https://www.dfrobot.com/wiki/index.php/Weather_Station_with_Anemometer/Wind_vane/Rain_bucket_SKU:SEN0186. [Accessed: 20- Jan- 2019].

[16] Sensirion, "Humidity and Temperature Sensor IC" Datasheet SHT1x V3.1, Dec. 2011 [Revised Dec. 2011].

[17] Bosch Sensortec, "Digital pressure sensor" Datasheet BMP180 V2.5, Apr. 2013 [Revised Apr. 2013].

[18] HopeRF Electronic, "Low Power Long Range Transceiver Module" Datasheet RFM95/96/97/98(W) V1.0, Mar. 2016 [Revised Mar. 2016].

[19] Espressif Systems, "Highly integrated Wi-Fi SoC" Datasheet ESP8266EX V6.0, Nov. 2018 [Revised Dec. 2015].

[20] S. S. Haykin, Neural Networks And Learning Machines. Delhi: Pearson, 2016.

[21] C. M. Bishop, Pattern recognition and machine learning. New Delhi: Springer, 2013.

[22] Corinna Cortes and Vladimir Vapnik, "Support-Vector Networks," Kluwer Academic Publishers, 1995.