Title: Development of application for cloud type classification using machine learning techniques

Research team: Teerach Noumanong , Ratchakrich Khamsemanan , Shane Ayron Thomas, Bhira Tayarangsee , Nopchapat Voraratchaikul
Project advisor: Mrs.Bannaruck Tanjaphatkul , Mr.Sorapong Somsorn
School : Varee Chiangmai School

Abstract

Weather greatly impacts daily life, influencing activities and global climate patterns. With climate change making weather more unpredictable, we face challenges such as storms, floods, water shortages, and landslides. These issues highlight the need for better weather monitoring systems. Clouds are essential in regulating temperature, driving the water cycle, and shaping weather. Since cloud types are directly linked to weather changes, identifying them helps predict short-term and long-term weather events, such as storms and extreme temperatures, improving preparedness. However, identifying cloud types requires expertise, and inconsistencies can affect the accuracy of weather predictions. This project aims to create a web-based application that uses machine learning to automatically classify cloud types. The goal is to make cloud identification more accessible, assisting in environmental monitoring and raising awareness of weather patterns. The project also promotes citizen science, encouraging public participation in improving weather forecasting and climate action.

In this project, a machine learning model was trained to classify eight cloud types using a dataset of images. The model was tested with 175, 200, and 250 images to examine how dataset size affects accuracy. The best results came from using 175, 200, and 250 images, achieving a training accuracy of 96.5% and validation accuracy of 82.8%. The developed app, "Qmulo," performed well in classifying clouds, with Cumulonimbus clouds identified with 90% accuracy, Altocumulus at 88.57%, and Cirrus at 85.7%. When focusing only on correct classifications, Cumulus clouds were identified with 98.52% accuracy. While the application performed well overall, some cloud types, like Cirrocumulus and Stratocumulus, were classified with lower accuracy. This can be improved by expanding the dataset and selecting more representative images. The app shows potential in environmental monitoring and encourages public involvement in weather tracking. In conclusion, the "Qmulo" application demonstrates the effectiveness of machine learning in cloud classification, with room for future improvement in accuracy and broader public participation.

Keywords: Machine learning Cloud Classification Weather monitoring Citizen Science

Acknowledgement

The Development of a Web Application for Cloud Type Classification Using Machine Learning Techniques, has been successfully completed as planned.

The research team would like to express our sincere gratitude to Varee Chiangmai School for providing support in terms of facilities and learning equipment throughout the research process. We extend our appreciation to Mrs. Bannaruck Tanjapatkul, the project advisor, and Mr. Sorapong Somsorn for his guidance on machine learning and application development. Our gratitude also goes to all the Science and Technology Department teachers for their valuable advice and support.

Furthermore, we would like to thank our friends and close acquaintances for their encouragement in every aspect, as well as our families for their unwavering support and motivation, which have been instrumental in the successful completion of this project. The research team would like to express our deepest appreciation on this occasion.

Research Team

List of contents

List of Figure

Introduction

Weather plays a crucial role in our daily lives, influencing everything from the activities we plan to the broader patterns of climate around the world. In recent years, the impact of climate change has led to more unstable weather, causing serious challenges like water shortages, storms, floods, and landslides. These challenges emphasize the urgent need for improved weather monitoring and prediction to reduce risks and enhance preparedness.

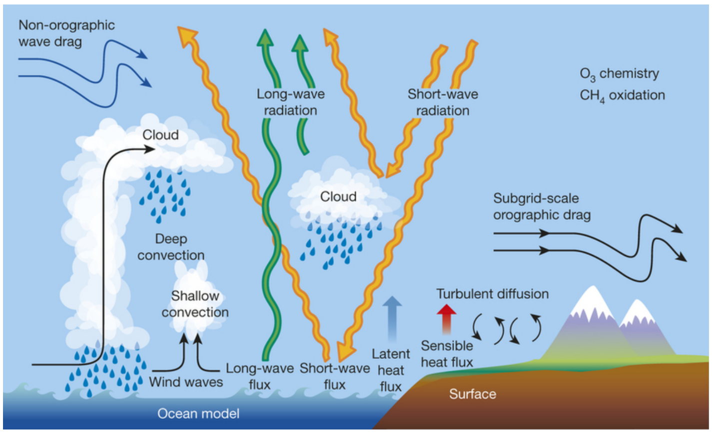


Figure 1: The Influence of Clouds on the Earth's Weather Patterns
(Source: https://shorturl.at/XLgig, 2024)

Clouds are a key component of the Earth's atmosphere. They play a vital role in regulating temperature, driving the water cycle, and shaping wind patterns. Because cloud formations are directly linked to changing weather conditions, identifying different cloud types can help predict both short-term and long-term weather changes. Recognizing cloud types allows us to detect approaching storms, extreme temperatures, and other hazardous weather events, helping individuals and communities plan and prepare. For instance, farmers can use this knowledge to make informed decisions about irrigation and harvesting, pilots can anticipate turbulence, and emergency responders can act more swiftly in the face of severe weather.

In addition, analyzing long-term cloud patterns contributes to climate research, helping scientists better understand and address global environmental challenges. Despite its importance, accurately identifying cloud types remains difficult. It requires expertise and experience, and human error often leads to inconsistencies, which limits its usefulness for weather prediction and climate studies. To address this issue, the goal of this project is to develop a web-based application that uses machine learning to automatically classify cloud types. By making cloud identification easier and more accessible, this tool will enable individuals to take part in environmental monitoring, raising public awareness of atmospheric changes. This initiative is in line with the principles of citizen science, encouraging collective efforts to improve weather forecasting and support sustainable climate action.

Objectives

1. To develop a web application for cloud classification using machine learning techniques.
2. To evaluate the effectiveness of the developed cloud classification tool.

Expected Outcomes

1. To enable the general public to easily classify cloud types.
2. To promote environmental monitoring as a fundamental aspect of citizen science.

Scope of Study

1. The study involves programming a machine learning model to learn and classify cloud types. Initially, the system will be trained to identify eight cloud types: Cirrus, Cirrocumulus, Cirrostratus, Altostratus, Altocumulus, Cumulus, Stratocumulus, and Cumulonimbus.
2. The developed tool will be implemented as a mobile-friendly web application to enhance usability.

Project Duration

October 2024 – February 2025

Literature Review

Cloud Formation and Types

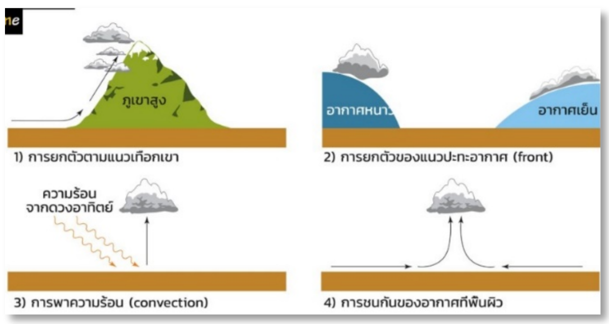
When solar radiation reaches the Earth, different surfaces absorb heat at varying rates. Water bodies, in particular, absorb heat and cause water to evaporate into water vapor, which then rises into the sky. As this vapor ascends and encounters cooler air in the upper atmosphere, it condenses into tiny water droplets that cluster together to form clouds. The natural process of cloud formation can be categorized into four main types:

- Orographic lifting – When air is forced to rise over a mountain range.
- Frontal lifting – When air masses collide, causing warm air to rise over cooler air.
- Solar heating – When heat from the sun causes water vapor to evaporate and rise.
- Convergence of air at the surface – When winds push air upward, lifting a mass of air. (Santi Phailoblee, 2023)

Due to these different formation processes, clouds appear in various shapes and structures. Some clouds form as thin, wispy layers when the air has low moisture content, while others may look like cotton-like puffs or delicate, feather-like streaks. Scientists classify clouds into two main groups based on their shape:

- Cumuliform clouds (Cumulus) – These clouds have a dense, fluffy appearance and are often separated by clear spaces in the sky.
- Stratiform clouds (Stratus) – These clouds form in thin, layered sheets and may cover the sky in uniform layers.

Both types of clouds can be found at different altitudes, which are divided into three levels: high-level, mid-level, and low-level clouds. Based on their shape and altitude, clouds are further classified into ten different genera, as shown in the following figure.



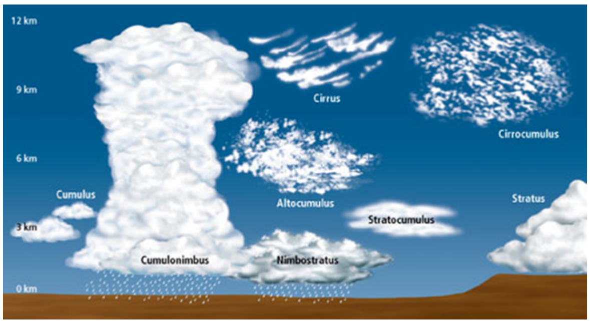


Figure 2: The Process of Cloud Formation. Figure 3: Types of Clouds

(Source: http://www.mitrearth.org/23-2-clouds-fog-dew/ , https://www.noaa.gov/jetstream/clouds/nws-cloud-chart, 2023)

Machine Learning Technique through CNN and Web Application Development

For the cloud classification web application developed in this study, the chosen machine learning technique is Convolutional Neural Networks (CNN). CNN is a type of artificial neural network designed specifically for processing and analyzing images. It works by breaking down an image into various components, such as lines, shapes, and patterns. These characteristics are then used to identify and classify objects within the image. CNNs are particularly effective for image recognition tasks because they can detect subtle features and patterns that may be difficult for humans to identify.

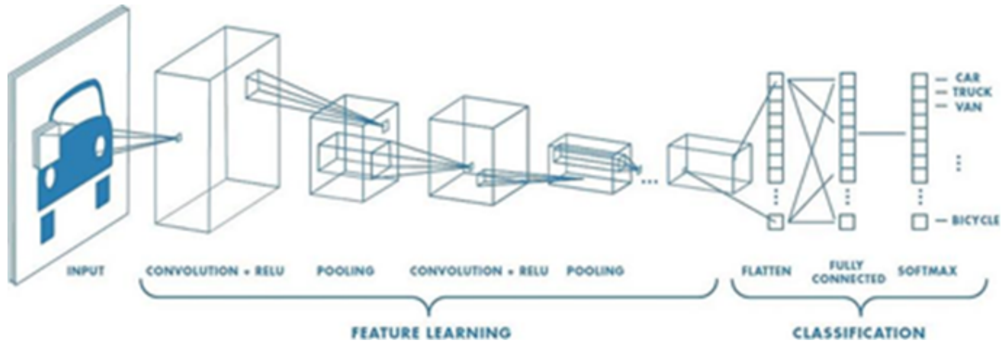


Figure 4: The Working Principle of CNN (Convolutional Neural Network)

After evaluating the performance of the cloud classification tool, the next step is to develop the application using "Streamlit", a Python-based framework that simplifies the creation and deployment of machine learning and data science applications. Streamlit enables developers to quickly build interactive web applications with minimal coding effort, making it a popular choice for rapid prototyping and deployment within the data science community.

Additionally, the application will be adapted for use on mobile devices using Android Studio, an integrated development environment (IDE) designed for building Android applications. Android Studio supports programming languages like Kotlin and Java, allowing developers to create high-performance, native mobile applications. By using both Streamlit for the web version and Android Studio for the mobile version, the application will be accessible on both desktop and mobile devices, providing a convenient way for users to classify clouds on the go. This combination of machine learning techniques and accessible development platforms ensures that the cloud classification tool is both powerful and user-friendly. It will offer a seamless experience, allowing users to easily classify clouds and participate in environmental monitoring.

Methods and Materials

Materials

1. Cloud Chart and Cloud Classification Guide (International Cloud Atlas)
2. Mobile Phone Camera
3. Kaggle Website
4. Google Colab
5. Streamlit
6. Android Studio

Methods

1. Collecting and Selecting Cloud Images
   - Cloud images were downloaded from Kaggle, a website with datasets for machine learning.
   - The images were checked manually to ensure they were clear and useful for training the program.
   - A total of 8,413 images from 5 datasets were reviewed, with 8 cloud types selected for training:

Cloud Type	Number of Images
Cirrus	51
Cirrostratus	51
Cirrocumulus	53
Cumulus	116
Altocumulus	96
Altostratus	51
Stratocumulus	50
Cumulonimbus	63

2. Training the AI Model
   - The program was trained to recognize cloud types using Google Colab.
   - The training was repeated 10 times with varying numbers of images.
3. Choosing the Best Model
   - The best model was selected based on its ability to classify clouds accurately.
   - The optimal results were achieved when training with 175, 200, and 250 images per set, achieving 96.5% accuracy during training and 82.8% accuracy in testing.

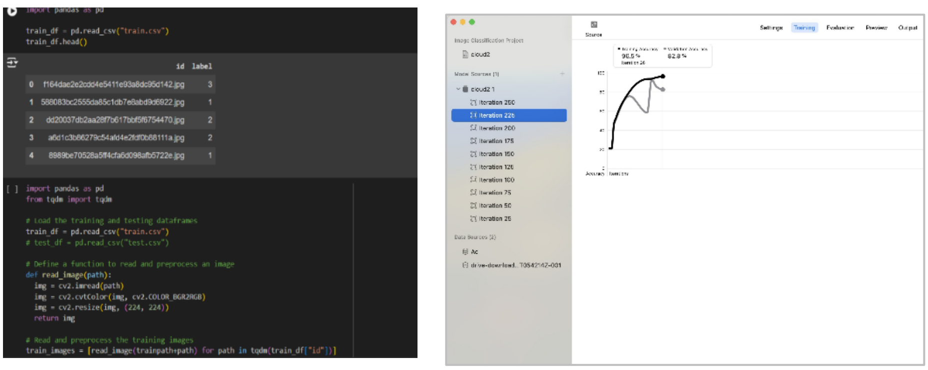


Figure 5: Machine learning procedure

4. Building the Cloud Classification Application
    - The best-performing model, trained with 250 images per set, was used to develop an application called “Qmulo” using Xcode.
5. Testing the App
    - The application was tested using a new set of cloud images to evaluate its performance.
6. Analyzing and Summarizing Results
    - The results were analyzed to determine the effectiveness of the app and the AI model in classifying cloud types.

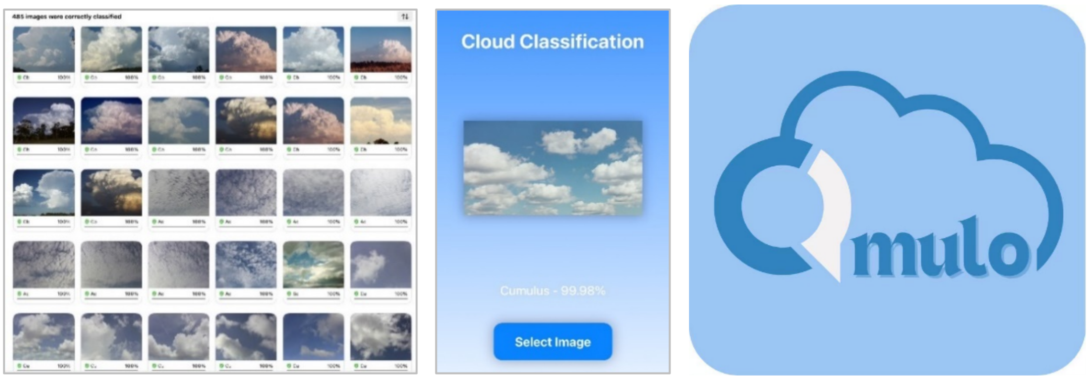


Figure 6: High-Accuracy cloud image dataset for developing “Qmulo” Application

Results and discussion

In this project, the machine learning model was trained to classify 8 types of cloud images. The goal was to investigate how the number of images in the training and validation sets affects the accuracy of the model. Different numbers of images were used in the training set—175 images, 200 images, and 250 images. The best results were achieved with 175, 200, and 250 images in the training set. With these sets, the model reached a training accuracy of 96.5% and a validation accuracy of 82.8%. This means that the model performed very well on the training images but was less accurate when it came to classifying new images that it had not seen before, which is why the validation accuracy was lower.

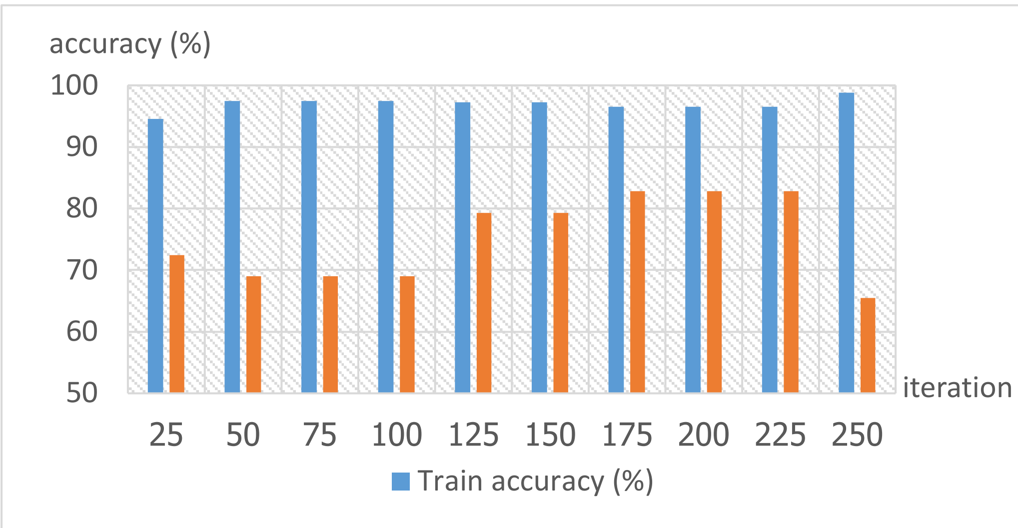


Figure 7: The results of machine learning training with different image training datasets

The graph shows how accuracy changes depending on the number of images used. It highlights how important the size of the dataset is in improving the model's ability to classify images accurately.

Then we used a training set of 250 images to create the cloud classified application called "Qmulo". The application was designed to identify and analyze different types of clouds, and it performed very well in doing so. Here’s a detailed look at how the app performed:

Accuracy Percentage in Testing: The application accuracy was measured by how well it identified different cloud types in the testing images. The best results came with Cumulonimbus clouds, which had the highest accuracy rate of 90%. This means the app was able to correctly identify Cumulonimbus clouds 90% of the time. Following this, Altocumulus clouds had an accuracy of 88.57%, and Cirrus clouds were correctly identified 85.70% of the time.

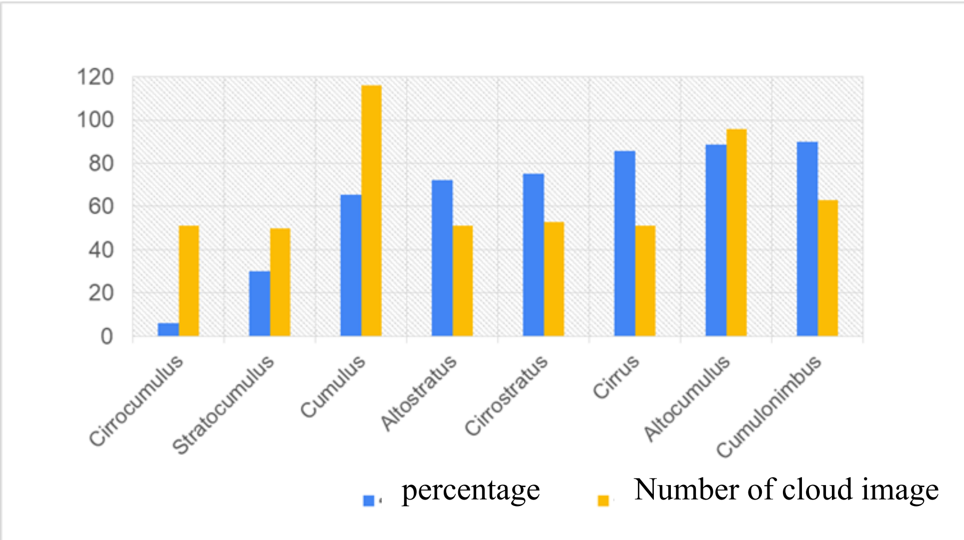


Figure 8: The accuracy of cloud type identification by the Qmulo application

Accuracy in Correctly Identifying Cloud Types: When only looking at the images where the cloud type was identified correctly, the application showed even better performance. All cloud types had an accuracy rate of over 85%. The Cumulus clouds were identified with the highest accuracy of 98.52%, meaning almost 99% of the time, the applicaiton correctly identified these clouds. Altocumulus clouds followed closely with an accuracy of 97.85%, and Cumulonimbus clouds had an accuracy of 96.27%. This shows that the application was especially good at recognizing certain cloud types, like Cumulus clouds.

Overall Performance: Overall, the app showed strong performance across all cloud types, with no cloud type falling below 85% accuracy. This means the application can reliably classify most clouds, but some types, like Cumulonimbus, might still need a little more training to improve its accuracy even further.

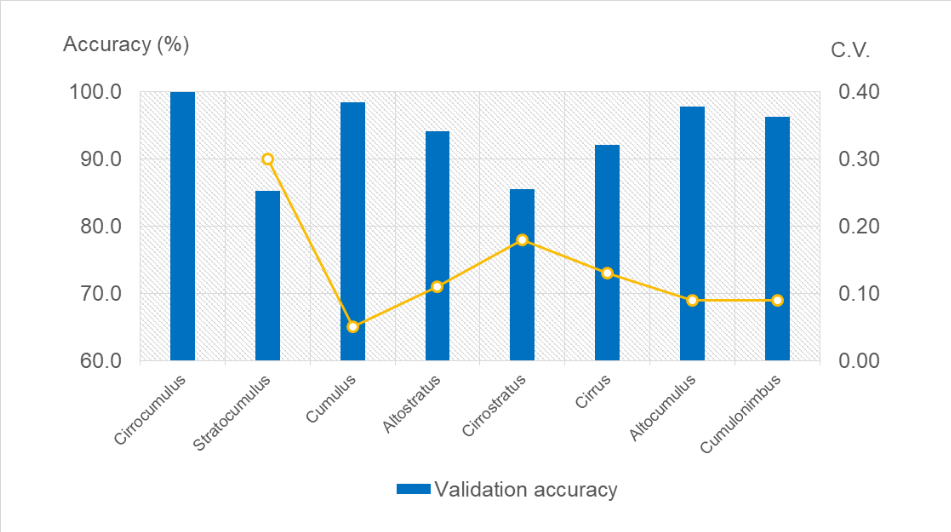


Figure 9: The accuracy of cloud type identification by the Qmulo application

Conclusion

Based on the performance test results for cloud type classification in the "Qmulo" application, the following summary categorizes the application's cloud identification capabilities into four levels:

Cloud Type	Correct-Incorrect ratio	Validation accuracy	C.V.
Cirrocumulus	5.88	100.00	(1 Picture)
Stratocumulus	30.00	85.21	0.30
Cumulus	65.38	98.52	0.05
Altostratus	72.41	94.18	0.11
Cirrostratus	75.00	85.53	0.18
Cirrus	85.70	92.09	0.13
Altocumulus	88.57	97.85	0.09
Cumulonimbus	90.00	96.27	0.09

Low performance:

This level refers to cloud types that the application identifies with less than 60% accuracy. Cloud types such as Cirrocumulus and Stratocumulus fall into this category. This low performance could be improved by increasing the dataset by at least 100 more images. Additionally, selecting images that clearly show these specific cloud types from an appropriate distance (not too close) would help the application better recognize them.

Moderate performance:

At this level, the application shows moderate accuracy in identifying cloud types, but with higher accuracy for certain types, like Cumulus clouds. The classification can be improved by selecting images that clearly display the characteristics of Cumulus clouds from an appropriate distance, ensuring the application has a better chance of identifying them correctly.

Good performance:

Here, the application demonstrates good accuracy in identifying most cloud types, but the variation in results is still somewhat noticeable. Cloud types such as Altostratus, Cirrostratus, and Cirrus clouds fall into this category. To improve classification, images of these clouds should be taken at an appropriate distance to better highlight their unique features, helping the application distinguish between them more easily.

Great performance:

This level indicates very good performance with high accuracy in both the classification percentage and the identification of cloud types, accompanied by a low coefficient of variation. Altocumulus and Cumulonimbus clouds belong to this category, where the app shows excellent accuracy in identifying these cloud types.

From the study, it is clear that using at least 100 images is essential to achieving good classification accuracy for all cloud types. The unique shapes and characteristics of certain clouds help improve classification accuracy, especially for those cloud types that look quite similar to others. Therefore, for cloud types that share similar features, having a wide variety of images in the dataset is crucial for training the app to perform better. The development team plans to continue improving the app’s performance, making it more practical and effective for everyday use. The goal is for the application to encourage people to observe and track weather patterns more closely, ultimately contributing to the growth of citizen science in the future.

Presentation clip

https://www.youtube.com/watch?v=7J2AuMJTwLE&feature=youtu.be

Citation

National Oceanic and Atmospheric Administration. (2023).NWS Cloud Chart. April 26 2024, available at https://www.noaa.gov/jetstream/clouds/nws-cloud-chart.

Methacharoenpha, Ratsarin. (2022). Application of Machine Learning to Industrial Work (Part 1). May 24 2024, available at https://www.nectec.or.th/smc/machine-learning/2567.

Institute for the Promotion of Teaching Science and Technology. (2022). ***Earth Science System.*** Bangkok: Intereducation Supplies.

Givemecode1655.(2021, January 12). Image Classification_with CNN :Teaching how to make CNN from start to finish, able to swat mosquitoes, available at https://youtu.be/JrtuWJYrmSM?si=3SpgFjWyx1vwPjcG.

GLOBE Data Entry (Atmosphere – Cloud Observation)

Varee Chiangmai School Weather, Latitude: 18.758349, Longitude: 99.015422
Elevation: 302.5 meters, SITE_ID: 180943

THE GLOBE PROGRAM การป้อนข้อมูลทางวิทยาศาสตร์ ยินดีต้อนรับ Bannaruck Tanjaphatkul

หน้าหลักของการป้อนข้อมูล / Varee Chiangmai school / Varee Chiangmai school weather / เมฆ

การสังเกตที่ผ่านมาสำหรับ เมฆ

From 2025-01-01 To 2025-03-04

	ทำการตรวจวัดที่ เวลามาตรฐานสากล	
1	2025-01-06 01:00 UTC	✖ ลบทิ้ง
2	2025-01-06 05:00 UTC	✖ ลบทิ้ง
3	2025-01-06 10:00 UTC	✖ ลบทิ้ง
4	2025-01-07 01:00 UTC	✖ ลบทิ้ง
5	2025-01-07 05:00 UTC	✖ ลบทิ้ง
6	2025-01-07 10:00 UTC	✖ ลบทิ้ง
7	2025-01-08 01:00 UTC	✖ ลบทิ้ง
8	2025-01-08 05:00 UTC	✖ ลบทิ้ง
9	2025-01-08 10:00 UTC	✖ ลบทิ้ง
10	2025-01-09 01:00 UTC	✖ ลบทิ้ง
11	2025-01-09 05:00 UTC	✖ ลบทิ้ง
12	2025-01-09 10:00 UTC	✖ ลบทิ้ง
13	2025-01-10 01:00 UTC	✖ ลบทิ้ง
14	2025-01-10 05:00 UTC	✖ ลบทิ้ง
15	2025-01-10 10:00 UTC	✖ ลบทิ้ง
16	2025-01-13 01:00 UTC	✖ ลบทิ้ง
17	2025-01-13 05:00 UTC	✖ ลบทิ้ง
18	2025-01-13 10:00 UTC	✖ ลบทิ้ง
19	2025-01-14 01:00 UTC	✖ ลบทิ้ง
20	2025-01-14 05:00 UTC	✖ ลบทิ้ง
21	2025-01-14 10:00 UTC	✖ ลบทิ้ง

Badges Descriptions

I AM A COLLABORATOR

Our project focuses on developing a web-based application called "Qmulo", which utilizes machine learning to automatically classify cloud types. With climate change making weather increasingly unpredictable, understanding cloud formations can significantly enhance weather forecasting and preparedness for extreme events like storms and floods. The project aims to make cloud identification more accessible, encourage public participation in environmental monitoring, and contribute to climate action.

Throughout the development of *Qmulo*, our team collaborated closely, with each member playing a crucial role in ensuring the project's success. Teerach and Bhira served as Project Managers & Data Collectors, leading the project, organizing tasks, meeting deadlines, and curating a comprehensive dataset of cloud images for model training. Shane, as the Machine Learning Specialist, developed and trained the model to classify cloud images into eight cloud types, fine-tuning the algorithm and analyzing how dataset size impacted classification accuracy. Nopchapat handled Data Preprocessing & Model Testing, ensuring that cloud images were properly formatted for training and evaluating model performance using different dataset sizes (175, 200, and 250 images). Teerach and Ratchakrich worked on App Development & User Interface, developing the *Qmulo* web application and making it user-friendly for public use. Through teamwork and dedication, we successfully integrated our expertise to create *Qmulo*, a tool that advances cloud classification and contributes to solving environmental problems using citizen Science.

I WORK WITH A STEM PROFESSIONAL

For our project, we had the great opportunity to work with Mr. Sorapong Somsorn, our technology teacher and project advisor. With years of experience and a history of guiding award-winning student projects, he gave us valuable advice that helped improve our research. He guided us in choosing the best machine learning algorithms for cloud classification and showed us how to fine-tune our model for better accuracy.

Mr. Somsorn also helped us prepare our cloud image dataset properly and solve issues related to dataset size and accuracy. Thanks to his support, we were able to make our *Qmulo* app more precise, especially in classifying tricky cloud types like Cumulonimbus and Cumulus. His advice made a big difference in improving our project and making our application more reliable. Beyond the technical help, Mr. Somsorn inspired us with his knowledge and dedication. He taught us important problem-solving and research skills, making this project a truly valuable learning experience. We are grateful for his support and guidance, which helped us grow as students and young researchers.