

INTRODUCE TO DATA SCIENCE

[FINAL TEAM PROJECT]

Prof. Anand Paul

양창엽 (2015118527), 허수호 (2013105119)

⌘ WHY YOU CHOSE THE TOPIC

There are many kinds of big data. Among them, sensor data are most widely used and are in the limelight in the fourth industry. Therefore, we decided to collect data on the environment (temperature, illumination, air quality, noise) at home through the Arduino to get indirect experience of the fourth industry. In addition to this, to give an example of the classification of big data, we have added a topic that can analyze the most popular car models in a given area by collecting car models and location data when a car is taken with a mobile application.

Collect	Describe	Analyze	Summarize
Arduino Sensor	iOS Application	Core ML	iOS Application

[Table 1] Project Probability theory & statistical methods

⌘ INTRODUCE TO PROJECT

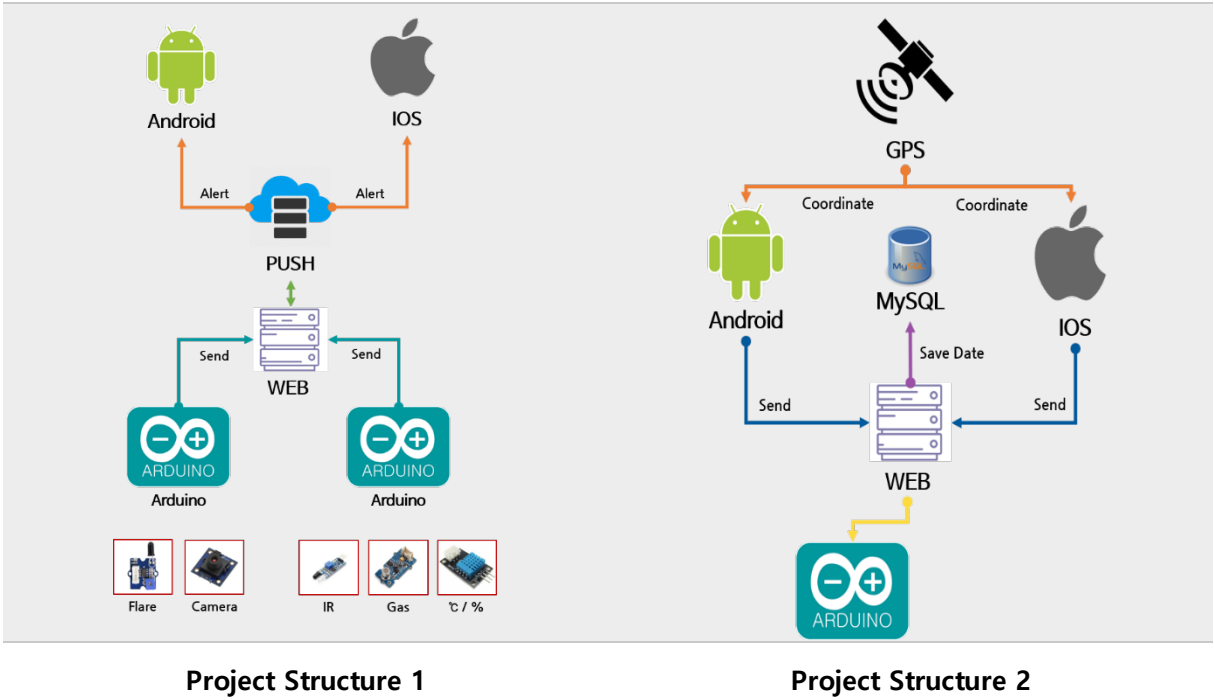
Our ASC team has been proceeding with two different projects.

- (1) The first project combines hardware. It uses the Arduino sensor to store data currently collected in the database and visually displays sensory data to the user through the iOS application and Android application by a map reduction process.

Number	Sensor Module Name
001	Flare Module Sensor
002	Gas Module Sensor
003	CDS Module Sensor
004	3 Colors LED Module Sensor (RGB)
005	Temperature Module Sensor
006	Humidity Module Sensor

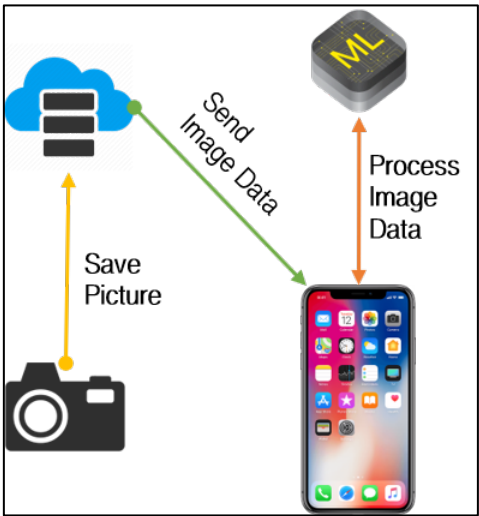
[Table 2] Arduino using Sensor Module list

Project Structure



[Table 3] Project Structure

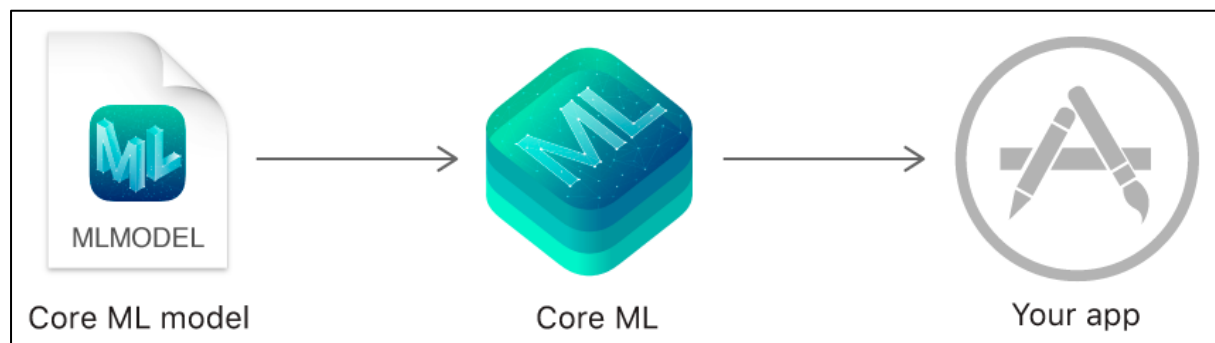
(2) The second is an advanced application project that scans the model of the car at the current location via Core ML which uses Apple Deep Learning technology. We analyze photos and classify car models and store them in the database. This is a project that shows the most notable models of cars around me through the Tableau graph.



[Figure 1] Classification Car Model Structure

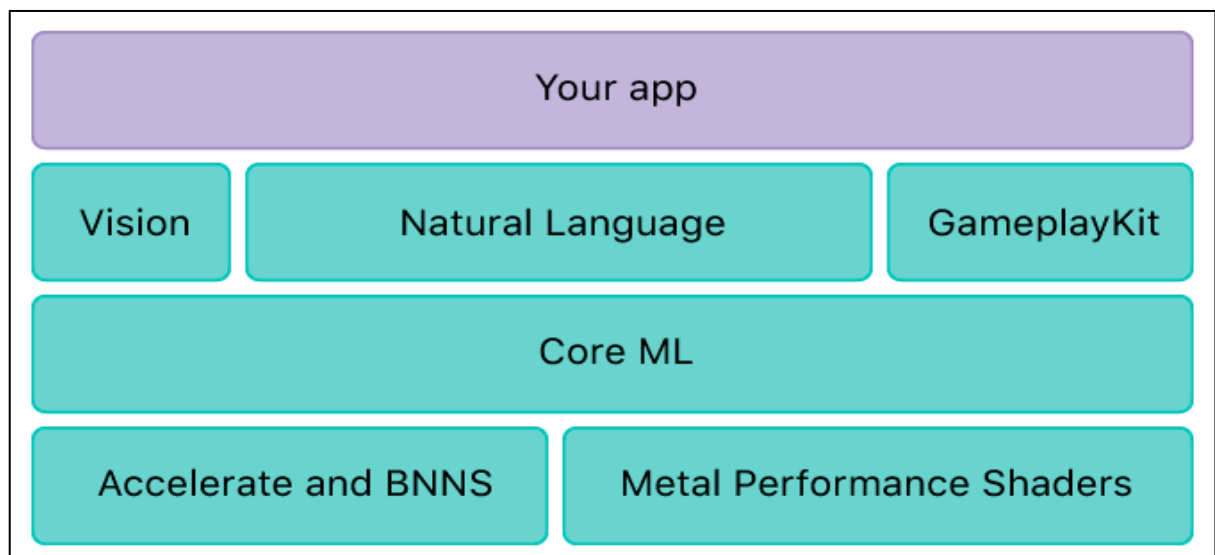
* WHAT IS CORE ML?

Core ML is the foundation for domain-specific frameworks and functionality. Core ML supports Vision for image analysis, Natural Language for natural language processing, and Gameplay Kit for evaluating learned decision trees. Core ML itself builds on top of low-level primitives like Accelerate and BNNS, as well as Metal Performance Shaders.



[Figure 2] Core ML Structure

Core ML is optimized for on-device performance, which minimizes memory footprint and power consumption. Running strictly on the device ensures the privacy of user data and guarantees that your app remains functional and responsive when a network connection is unavailable.

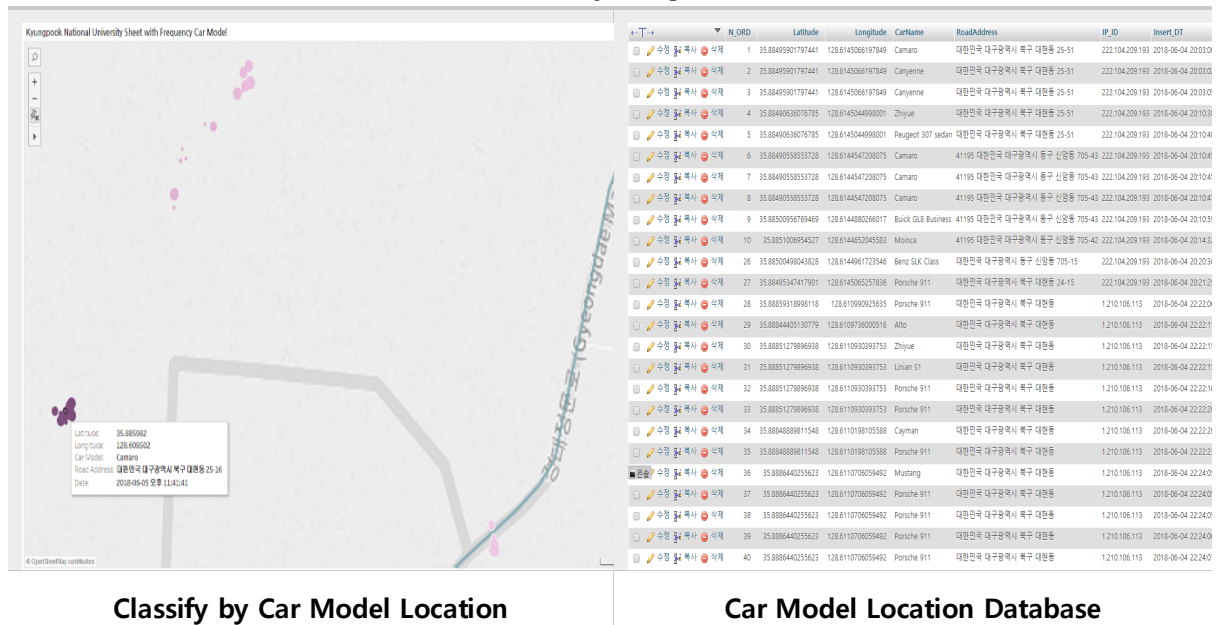


[Figure 3] Core ML App Structure

PROJECT DATA ANALYZE

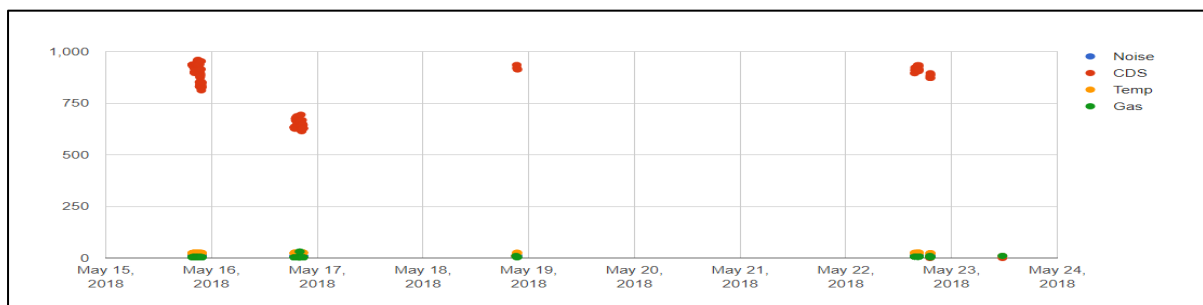
Using the Tableau data analysis tool, you can display the vehicle model names around Kyungbook National University and the time taken on the map screen through the coordinate value of dimensions. In addition, using the Google Fusion Table, we could see the map through the coordinate value of dimensions, and by comparing to Tableau, we could see more detailed data.

Tableau GEO Visually Image and Database



[Table 4] Tableau Geo Visually Image and Database

Analyzing sensory data (temperature, humidity, gas, CDS, noise) collected through Arduino using the Tableau data analysis tool, detailed information about the sensor around Arduino can be used to help manage the equipment based on the sensory data, like adjusting the temperature or lightening lights.



[Figure 4] Google Fusion Table by Arduino Sensor Dot Chart

Tableau Arduino Sensor Data by Visually Graph and Database

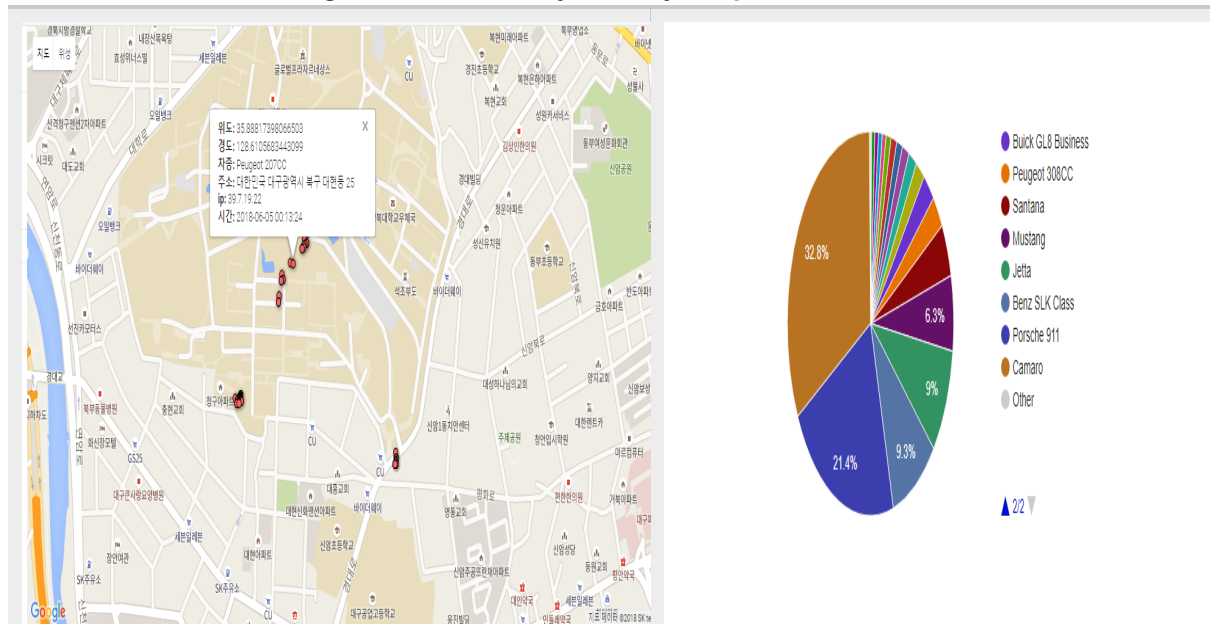


Analyze Arduino sensor data

Arduino sensor database

[Table 5] Tableau Arduino Sensor Data by Visually Graph and Database

Google Fusion Table by Visually Graph and Database



Classify by Car Model Location

Classify by Car Model Pie Chart

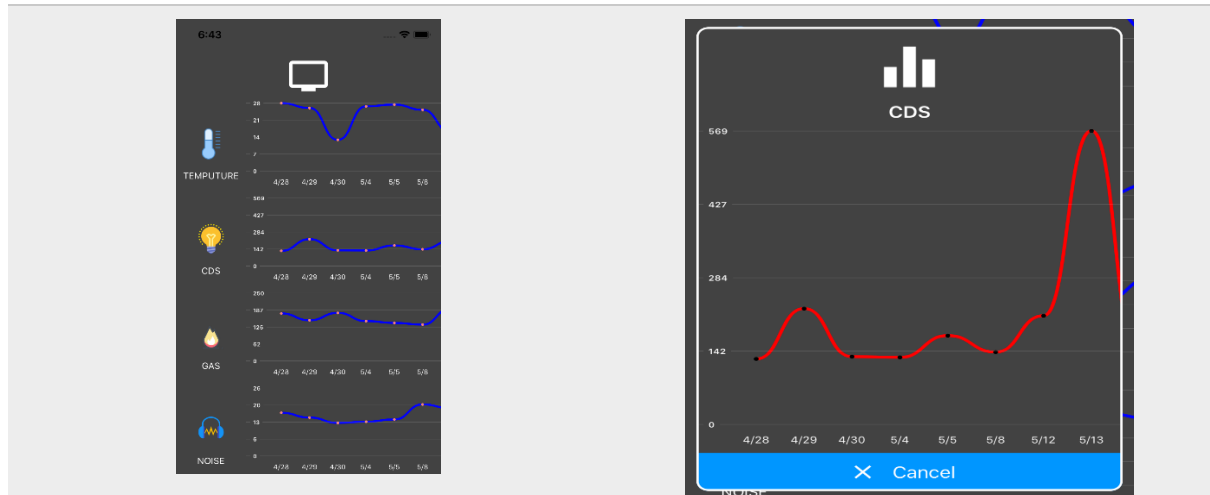
[Table 6] Google Fusion Table by Visually Graph and Database

PROJECT DRAFT

Project Source and Data Set Link:

<https://drive.google.com/drive/folders/1bHfc66WfkB1YLNifW3r4q2O7ZLe179Qb?usp=sharing>

Visually Arduino Sensor Data Graph Display



[Table 7] Visually Arduino Sensor Data Graph Display



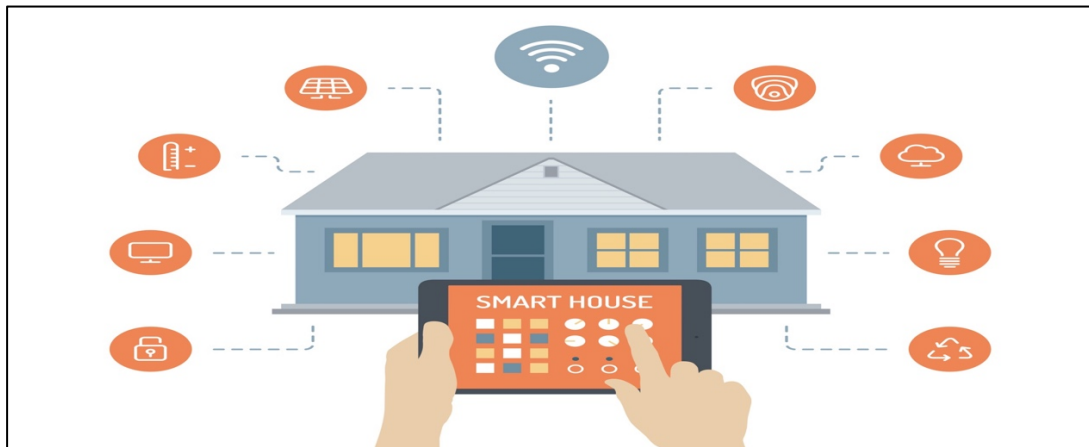
[Figure 5] Range Car Demo Play (Click Image)

Application Method List

- ⌘ The weather station will ask you for current weather information.
- ⌘ For clear image quality, we request the concentration of fine dust at the current position from the weather agency.
- ⌘ Analyze the model of the car through Core ML.

PROJECT EXPECTATION EFFECTIVENESS

- Data collected through Arduino can be analyzed various weather-related information in the current location.
- Arduino can be used to build home IoT systems and car analysis models at low cost.
- It is possible to analyze the car image obtained from CCTV and to prevent crime on getaway vehicles and drunk-driven vehicles.
- The analyzed data of the car models can be used to obtain information about the surrounding car models and to study popular car models in sales.
- Based on the collected data from the Arduino, the home IOT system can be built.



[Figure 6] IoT Smart Home

FHWA Vehicle Classifications			
1. Motorcycles 2 axles, 2 or 3 tires 	2. Passenger Cars 2 axles, can have 1- or 2-axle trailers 	3. Pickups, Panels, Vans 2 axles, 3-tire single units Can have 1 or 2 axle trailers 	4. Buses 2 or 3 axles, full length
5. Single Unit 2-Axle Trucks 2 axles, 6 tires (dual rear tires), single-unit 	6. Single Unit 3-Axle Trucks 3 axles, single unit 	7. Single Unit 4 or More-Axle Trucks 4 or more axles, single unit 	8. Single Trailer 3- or 4-Axle Trucks 3 or 4 axles, single trailer
9. Single Trailer 5-Axle Trucks 5 axles, single trailer 	10. Single Trailer 6 or More-Axle Trucks 6 or more axles, single trailer 	12. Multi-Trailer 6-Axle Trucks 6 axles, multiple trailers 	
11. Multi-Trailer 5 or Less-Axle Trucks 5 or less axles, multiple trailers 	13. Multi-Trailer 7 or More-Axle Trucks 7 or more axles, multiple trailers 		

[Figure 7] Classify Car Model

1. Project Development Tool Environment

- 📁 XCODE Version 9.4
- 📁 Tableau Desktop Professional Edition 2018.1.1 64bit
- 📁 Apple Swift Version 4.1.2
- 📁 macOS High Sierra Version 10.13.5
- 📁 Google Fusion Tables Version Beta

2. Project Reference URL

- 📁 Wikipedia, Apache Hadoop,
https://ko.wikipedia.org/wiki/%EC%95%84%ED%8C%8C%EC%B9%98_%ED%95%98%EB%91%A1
- 📁 Apple Developer, Core ML, <https://developer.apple.com/documentation/coreml>
- 📁 Apache Hadoop, Hadoop, <http://hadoop.apache.org/>
- 📁 YouTube, Big Data Analytics,
<https://www.youtube.com/playlist?list=PLd1DDyoEDVmbvNUf6m6VjKh6KUqQt2UDd>
- 📁 Google Drive, Google Fusion Tables, <https://sites.google.com/site/fusiontablestalks/stories>
- 📁 gifs, iOS APP Operation GIF, <https://gifs.com/gif/ios-voMgvX>