



MACQUARIE
University
SYDNEY • AUSTRALIA

Report on “Detect Drive off – An Automated Detection, Payment, and Reporting System.”

Name: Karthik Varma Keerthipati

Student ID: 46272496

Unit Code: COMP8230

Assessment Task - 1

Session 1 - 2021

Submitted on 26 March 2021

Contact: karthikvarma.keerthipati@students.mq.edu.au

Table of Contents

1. Executive Summary	3
2. Industry Context – Domain	Error! Bookmark not defined.
2.1 Introduction	4.
2.2 Data Islands.....	Error! Bookmark not defined.
3. Motivating Scenario.....	Error! Bookmark not defined.
4. Problem Statement.....	Error! Bookmark not defined.
5. Approach	Error! Bookmark not defined.
6. Results/Prototypes.....	Error! Bookmark not defined.
7. Conclusion	Error! Bookmark not defined.
8. References	Error! Bookmark not defined.

Figures

1. Capturing the image and detection of vehicle.....	6
2. Life cycle of Data Mining Model.....	7
3. Flowchat of Image capturing and Identification.....	8.
4. Knowledge Lake with ML services.....	9
5. Prototype for Model	10

1. Executive Summary

The technologies are upgrading day by day but in some areas still using traditional process only, even though there is a scope to utilize the latest emerging technologies. The drive offs in service stations are one of the places where we are focusing now to reduce the crime. These are happening regularly in most of the locations intentionally or fail to pay for the fuel. We can stop this by using the Data mining, machine learning, IOT devices, Hadoop, and artificial intelligence.

This will also be helpful to police in identifying the suspects and investigation purposes. The main idea is to use the vehicle registration number to identify the vehicle with time and location attached as well and should be stored in the centralized database to get access to all stations if the vehicle did the same thing again in another station, it would alert that owner and tagged that vehicle as drive off and will be removed that tag once the payment has done.

In any crime, mostly vehicles were used. So, with this system it is very useful to trace the suspects based on the movements and informs it to the police if the customer is having repeated drive offs or any crime history in the database.

The drive off tag should be allocated to that vehicle from the vehicle number recognition system and the payment notice should be sent automatically to the mail address of the customer from the RMS database taken as the distributed system to check and verify the customer. The customer must clear the payment within the due date, then the flag will be removed otherwise the report will be registered automatically and sent from the police.

There is no need to process the continuous stream of data from the cameras all the time, we can take pictures of the vehicle once the pump has started filling and the picture is stored with time and location as well. The transaction has not been completed within the allocated time, then the recognition system will start and performs the identification and payment process if there are multiple previous drive offs.

It will alert the staff and if the previous payment is due in any other station before, then the access from the pump should be declined and the information with location will be sent to the owner and police. The money will be recovered easily and reduces the time and effort of human intervention in dealing these cases.

The police can also easily track wanted persons with the help of our system because any vehicle must have to fill fuel at the service stations. Our system will alert the stores spontaneously once the wanted number on the plate has been recognized and report to the police. This is an added advantage with this system.

2. Industry Context - Domain

2.1 Introduction to Crime Investigation:

The crime investigation is the process of finding and collecting the clues or evidence for a case to identify the culprit. This involves activities such as information analysis, and critical observational skills. There are several types in criminal investigation such as fraud investigation, criminal defense investigation, kidnapping investigation, and homicide investigation. The theft investigation is also a type of criminal investigation.

The increase in crime rate will create the new types of crime investigations. There are mainly four steps in investigations. They are Examine, Photograph, Sketch, and Process. We are going to use these steps in the below research.

There are some traditional and time-consuming methods in investigating the crime scenes. But with the arrival of latest techniques and technologies such as IOT devices and machine learning models, we can easily investigate and get the accurate results.

There are lot of improvements in the technologies that are helping the investigations rapidly. The data is enormous to investigate with lot of social data from various sources. We need big data technologies and data mining tools to maximize the use of data and to derive insights from the data.

This will give the better understanding on the case and provide some useful information on the scene. The crime rate will be decreased by alerting the people in advance and can protect the society without any harm.

2.2 Data Islands:

The automation of the detection of drive offs needs to be monitored from various devices and the data from various sources might be useful for the detection. The data islands are explained below.

a. IOT Sensor data: The images form the CCTV camera and other surveillance cameras, the location from the location sensor.

b. Social data: The data from tweets, messages, text, and mentions will be used to improve the performance of model and investigation.

c. Private data: The data from IoT devices, and networking devices to capture and process the data.

d. Historical data from centralized database to check the history and RMS database to get the information of the vehicle owner.

In future if there is new type of data needs to be collected then the related data island will also be added.

3. Motivating Scenario

The crime investigation is the area where there is huge necessity for data mining because of huge data for various sources having structured, semi-structured and unstructured data. It is very hard to store and process the huge amounts of different kinds of data. So, big data technologies will be used to solve the problem. The millions of dollars' worth of fuel is recorded as fail to pay from the service stations. Even though, there is a lot of CCTV cameras working still it is not possible to stop and control the activities with the help of continuous observation from a human.

There should be automated system which tell us the details of the vehicle before giving access to the pump and there is also another chance to identify the suspected criminals with the help of our system if worked together with the police without adding any extra efforts to the system. Big data technologies are playing a very important role in this research to deal with the data and in deriving insights as well.

The theft investigation is the important branch, we need to focus and made some research on this field. The rise of motor vehicle thefts every year needs to be controlled and the main idea behind this research is to stop the fuel thefts, identify the stolen vehicles and to inform the police if anything suspicious. Any vehicle must reach the service station to fill the fuel, so we can easily identify such vehicles if we got the information from the police into our systems and send the information back when matched.

The applications involved in crime investigation are image collection, evidence mining, forensic analysis, network analysis, web analysis, clustering techniques, and Multi-layer perceptron (MLP) models. It will provide great insights in the industry by recovering lot of money and speed up the investigation.



4. Problem statement

The 50 million dollars' worth of fuel is being stolen every year in all over Australia, which is most of them are using with the stolen cars and using these cars for other crimes as well. There are 56,300 motor vehicles were stolen in a year at a rate of about 154 thefts a day. After considering the problem seriously, we can observe that the theft percentage is increasing every year and with the normal process it usually takes very long to identify the offender by watching the CCTV footage to identify the person and reporting to the police. In this way, we cannot even know that the customer had already did any fuel drive offs or not. The customers who are performing regular fuel drive offs may not be identified with lack of centralized communication with all the other stations. The companies are losing huge amounts and not able to recover most of the money. It is also a challenge to trace down the offenders. The fail to pay is also a crime whether it is intentionally or not.

The monitoring of each vehicle all the time might not be helpful with human observation because every station might not be able to afford the technology involved. So, the proposed improvement over the existing one is to use the number plate recognition system which is very well known now a days. The number should be extracted, processed, and identified from the database to assess the previous drive offs or crime history and to finalize access to the customer.



Fig 1: Capturing the image and detection of vehicle.

It is also not that hard to identify the number plate because the vehicle is stationary at that time. The use of modern technologies such as machine learning and data mining, we can easily achieve the desired results. We can link the number obtained to the police database with time and location to ease the reporting process if the customer has not done the payment even after sending the payment notice. The police can also utilize this system by uploading the details in the database to trace the criminals and offenders when they drove into the station to fill the fuel.

5. Approach

Detect Drive Off is an application prototype to help and transform the traditional way of finding and reporting the fuel theft in all the service stations. This prototype will also help to identify the stolen cars and criminals who are in search of police.

The research is followed by using the CRISP-DM methods. Cross Industry Standard Process for Data Mining model contains six phases of life cycle.

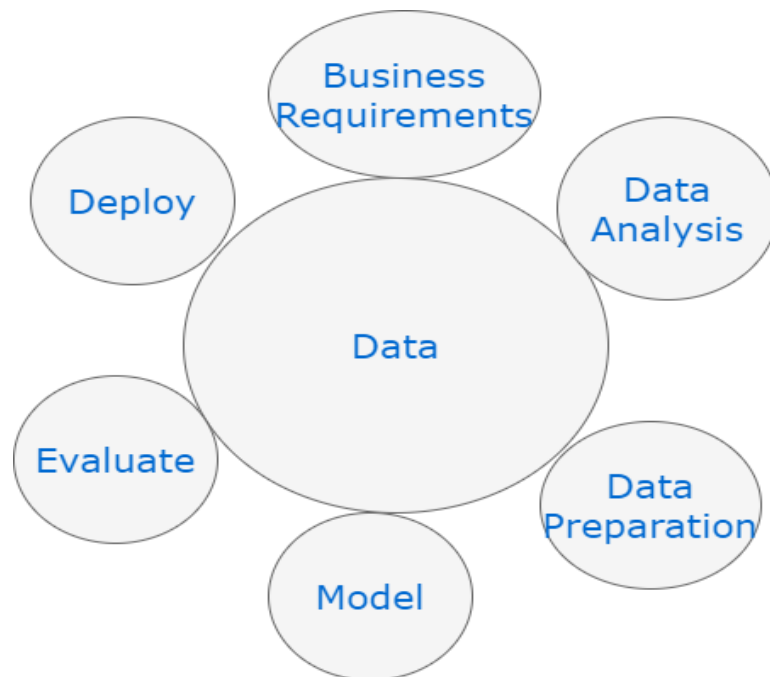


Fig 2: Life cycle of Data Mining Model.

There are two types of components where detection and identification is the first step and alerting, informing or reporting is the second step.

The image is captured from the camera once the pump has started. The system analyzes the vehicle number of from the image using character recognition technique and communicates with the databases for previous driveoff or criminal history and alerts the owner of that store and if anything is suspicious, it will automatically starts recording and inform to the police with the time and location of the store. The police are looking for some stolen vehicles, if they uploaded the details in the database the system will automatically flags the vehicles for lookup and search for that vehicles in the database to know the information like when, where and what time vehicle filled fuel and can easily identify on the spot if they are filling and in real time inform to the police. It will also very helpful to the police for investigation.

Data Ingestion, Cleaning, and Transformation:

The private data will be collected from the service stations as images which is unstructured data. We are using some techniques like data mining, text analytics, and natural language processing to find patterns in the data. Data Ingestion is the step to ingest the private data into the distributed file system in order to process the big data. There is huge amounts of data to process, so we need big data technologies. The data needs to be cleaned by removing some noise in the images. We can use the vision technology provided by various services like Google, IBM, Apple, and Amazon to reduce the work load and improve the efficiency.

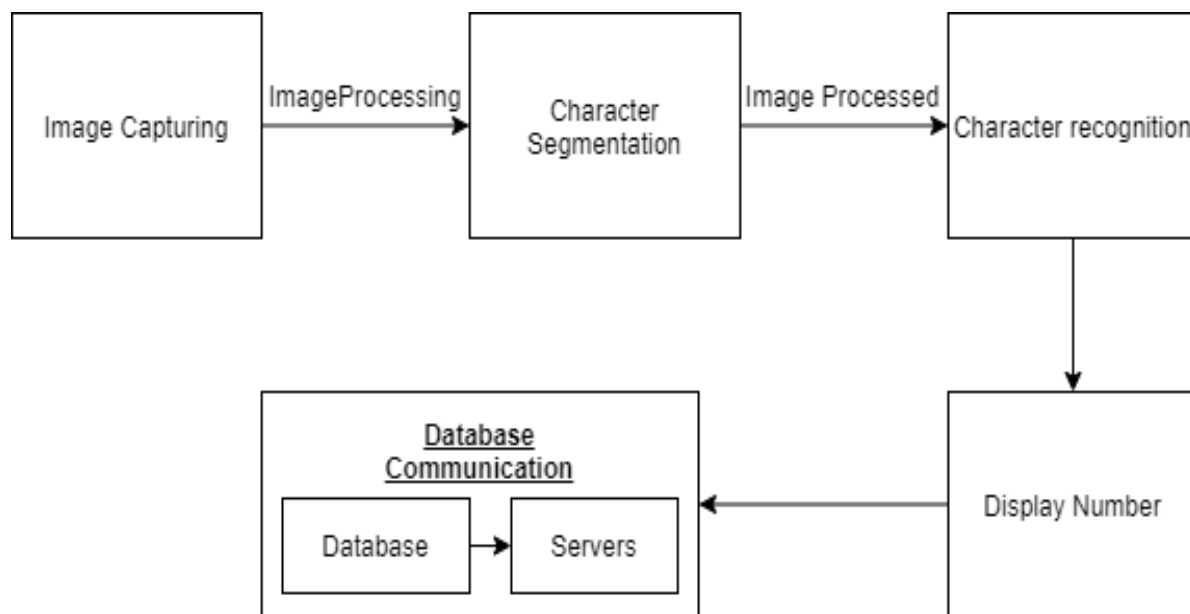
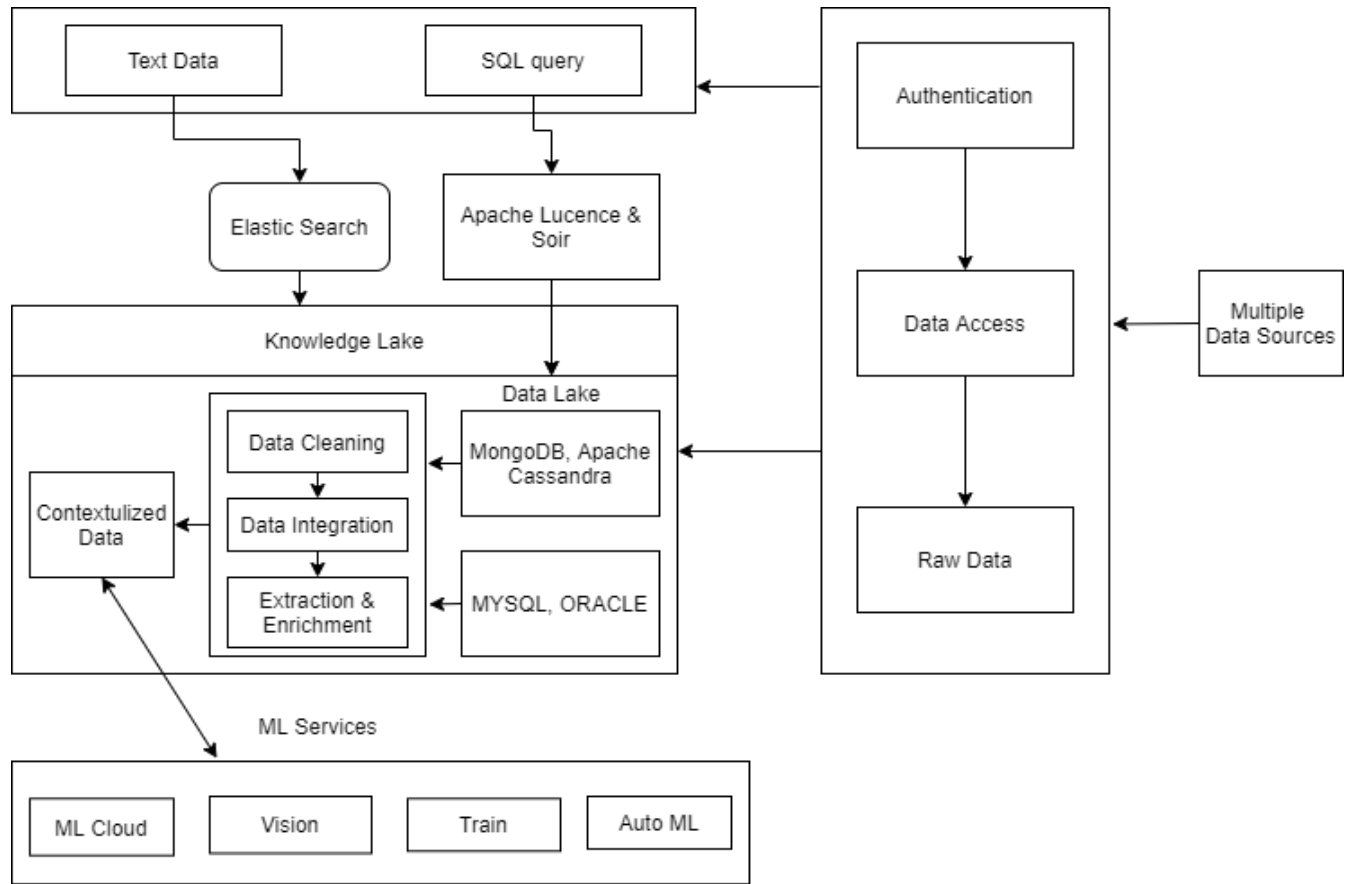


Fig 3: Flowchat of Image capturing and Identification

The character recognition can be done by detecting the text and splitting the images of characters and validating by confidence.

The Knowledge Lake Architecture:

The raw data from various locations are ingested to the database and processed. The data will be recognized and accessed through secure systems for authorization and authentication. The elastic search is used for real time data with fast search by indexing. The data is tagged if found any and processed for data visualization to get the insights from the data. The information is stored in the database after cleaning the raw data and search for incident history and looking for matches in the suspected list from the police database. The ML cloud is used for better predictions and storage with supervised machine learning techniques.



Knowledge Lake with ML Services

Fig 4: Knowledge Lake with ML services

We are using knowledge base for the analytics and identification of the vehicle. The ML services contains Keras, Sklearn, Numpy, and Core ML to run the model more accurately and effectively. In order to get the accurate result, we are building and training the model with deep neural networks. Google vision can also used to obtain better results in optical character recognition. The useful data can be stored in the knowledge base to get use of it when ever needed. The ML cloud, Vision, Train, and Auto ML speed up the process and produce the results effectively.

6. Results/Prototypes

The Detect Drive Off prototype is designed to automate the existing methods and to reduce or stop the crime by helping the police in the investigation. The prototype model is shown below.

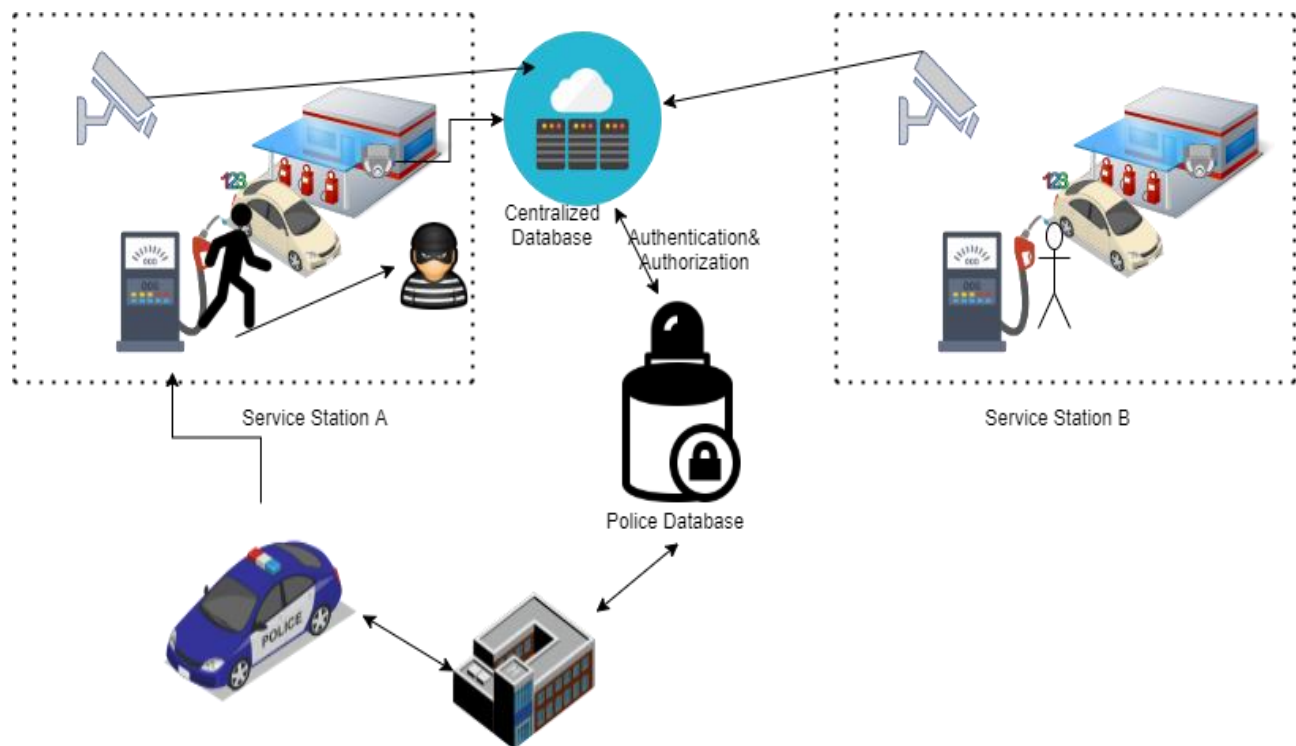


Fig: 5 Prototype for model

In the prototype, the camera took a picture of a car when he started filling and recognizes the vehicle number as result by evaluating the confidence of characters. It communicates with the centralized server to check the previous drive offs or any existing drive off payment dues. If there is anything suspicious in the record, it will alert the owner. If police are looking for someone or any vehicle, they can also upload those vehicle numbers in the database once it is searching for the history will also check those details and found any match immediately send the information automatically to the police and helps for the investigation. We can validate our results with the already detected suspects. This system is having a lot of scope in other fields as well because there is a huge increase in criminal activities every year and this system will help to reduce those activities by alerting and supporting the investigation to the police.

7. Conclusion

In this research, the most significant data mining techniques were used for crime investigation. We can also hope to integrate this approach in the enterprise editions to the fuel retailers. This will save a lot of money to the fuel retailers and can help the police in crime investigation.

The use of the detection and reporting system is to identify and report to police and the main purpose of this research is to solve the crimes fast. Our supervised learning model will give the accurate results. The system is developed in a way to implement and embedded with future technologies to process faster. The multiple articles and research paper on related technologies were considered and assessed to implement on this system.

A safer environment is to be maintained and given priority in the society to maintain peace. This system will try to stop most of the fail to pay incidents and reducing losses to the retail owners. The usage of vehicles is also increasing daily which we need to be focused to limit the usage.

The technology can be used for good purpose only and also not to disturb or violate the other circumstances.

8. References

(2021). Retrieved from Elastic Search: <https://www.elastic.co/what-is/elasticsearch>

Behesti, A. (n.d.). *Knowledge Lake Architecture*. Sydney.

data science pm. (2021). Retrieved from CRISP DM: <https://www.datascience-pm.com/crisp-dm-2/>

Internet of Things (IOT) Security. (n.d.). Retrieved from Imperva:
<https://www.imperva.com/learn/application-security/iot-internet-of-things-security/>

Lake Illawarra Drive off. (n.d.). Retrieved from Drive OFF Lake Illawarra:
<https://www.illawarramercury.com.au/story/6619898/heres-how-many-people-drove-off-without-paying-for-their-fuel-in-lake-illawarra/>

License Plate Recognition. (2021). Retrieved from Medium: <https://medium.com/programming-fever/license-plate-recognition-using-opencv-python-7611f85cdd6c>

Machine Learning Models. (2020). Retrieved from Analytics Vidya:
<https://www.analyticsvidhya.com/blog/2017/09/common-machine-learning-algorithms/>

NSW.gov.au. (2021). *Police NSW*. Retrieved from Police NSW:

https://www.police.nsw.gov.au/crime/frauds_and_scams/fraud_categories/fail_to_pay_for_fuel#:~:text=The%20offence%20of%20Fail%20to,station%20industry%20to%20the%20police.

Oracle. (2020). *Crime pattern Detection*. Retrieved from Crime Pattern Detection:

<http://cs.brown.edu/courses/csci2950-t/crime.pdf>

The cars to be stolen mostly. (2020). Retrieved from Sydney Morning Herald:

<https://www.smh.com.au/national/nsw/revealed-the-cars-most-likely-to-be-stolen-and-from-where-they-are-being-taken-20201215-p56nm7.html#:~:text=Nationally%2C%2056%2C300%20motor%20vehicles%20were,about%20154%20thefts%20a%20day.&text=Motor%20vehicle%20theft>

(Internet of things (IOT) Security, n.d.)