

School of Public Administration  
Bachelor of Science in Computing

**COMP491 Final Year Project  
Project Proposal**Academic Year 2018/19

|  |  |
| --- | --- |
| Pattern Recognition using Machine Learning | |
|  |  |
| Project number: | 3 |
| Student ID: | P-15-0792-1 |
| Student Name: | Zhiming Lin, Tony |
|  |  |
| Supervisor: | Dr. Yue Liu, June |
| Assessor: | Dr. Yang Xu |
|  |  |
| Submission Date: | 12th September 2018 |

Table of Contents

[1 Project Description 2](#_Toc524895021)

[2 Summary of Related Work and Key References 3](#_Toc524895022)

[2.1 Dataset preparation 3](#_Toc524895023)

[2.2 Others’ algorithms 3](#_Toc524895024)

[2.3 TensorFlow [4] 3](#_Toc524895025)

[3 Project Work Plan 4](#_Toc524895026)

[4 Risk Assessment 5](#_Toc524895027)

[References 6](#_Toc524895028)

# Project Description

Nowadays, the traffic congestion is one of the main problems in metropolitan cities, such as Macao. Many people complain about the time they have to spend on their commuting especially during the rush hours. It is inevitable that a solution should be came up to deal with the increasing traffic and pedestrians and worse-than-ever congestion on the road. In order to improve traffic efficiency, machine learning technology can help people to find the traffic patterns and predict the traffic conditions to avoid the traffic jams and reduce the volume of cars on the road all at the same time. It will be much more convenient for commuters to be aware of the precise real-time traffic information.

The aim of this project is to use machine learning technique to recognize the serious level of traffic congestion. The project is comprised of the following objectives:

* Identify the problem of image recognition: traffic pattern recognition.
* Do literatures review on image recognition and machine learning algorithms.
* Implement a machine-learning algorithm to recognise the traffic condition.
* Test the algorithm using real-time traffic image crawled from Transport Bureau of the Macao S.A.R. website (DSAT) [1] website and analyse the traffic results.
* Compose a final project report.

In this project, machine learning is used to analyse the traffic congestion by using instant traffic status images provided by DSAT. Users are able to understand congestion level of roads from the results of machine learning and get real-time information about specific roads.

During the development of this project, the data from DSAT may be tricky to craw. Moreover, image cleansing may be needed to perform since it is likely for computers to recognize and further analyse. Also, the outcomes of training data may not be good enough due to the camera limitations, weather conditions and complicated road environments.

# Summary of Related Work and Key References

## Dataset preparation

Dataset is one of the most important part of this project. Therefore, it should be considered seriously. There are two approaches can be conducted. The first is to find if there are suitable traffic images which can fit the situation of Macao perfectly. The second method is to make dataset which contains traffic images from DSAT. This method requires more time compared to the first one. Also, according to this paper [2], label method of vehicles should be stronger and well-designed network should be explored.

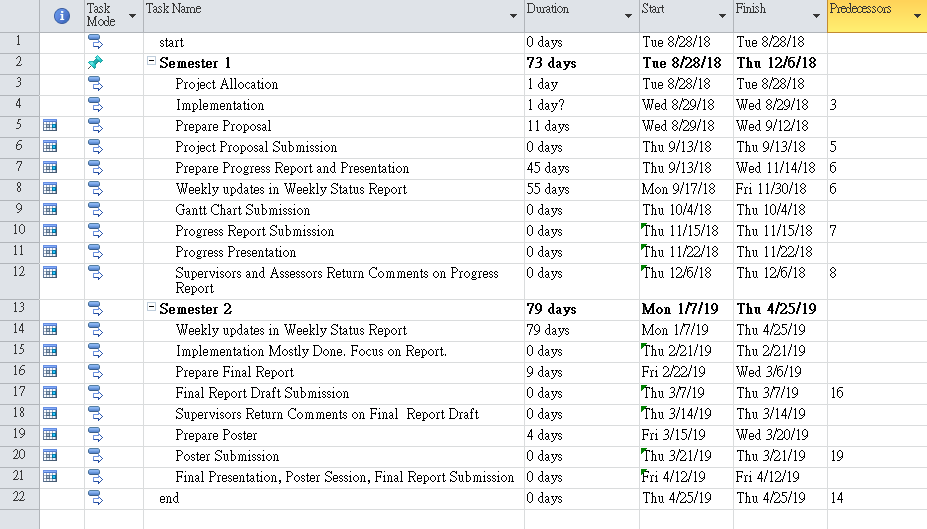
## Others’ algorithms

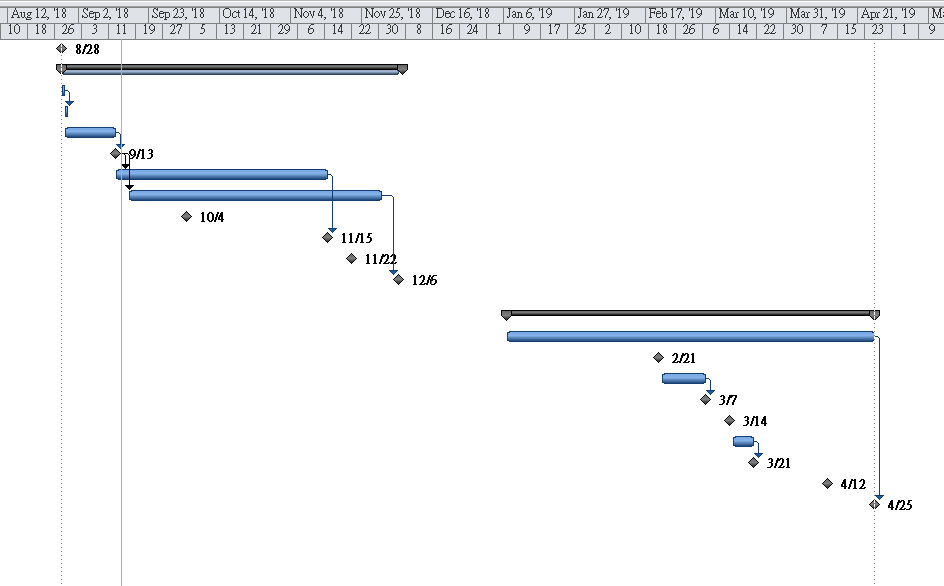
Nowadays, more and more people are coming to travel or work in Macao, and it is inevitable that this situation could worsen the traffic congestion in Macao. People are more likely to avoid wasting time on commuting. As this paper stated [3], the traffic camera is used to detect traffic intensity. They proposed a method to select a specific region manually from the images of traffic and performed filtering and attribute extraction operations to determine the road occupancy. However, this may not apply to Macao situation perfectly because the roads in Macao are quite narrow and the congestion level is better defined by machine automatically.

## TensorFlow [4]

TensorFlow is the tool which is going to be used in this project since it provides various of tools that can ease development. This tool is developed by Google which is the perfect tool for building deep learning network.

# Project Work Plan





# Risk Assessment

Risk one: training data may be lost.

Solution: It is important to have back-up training data because this is crucial to whole project. Thus, the data should have at least one copy in different places such as USB (Universal Serial Bus) flash drive or cloud drive. Once the original data is lost, the backup data can be used immediately.

Risk two: images quality may be too low to train.

Solution: The image quality is another significant point which affects the quality of the training data and the accuracy of the training results. In order to prevent such things happening, images which are from DSAT website when the weather is raining or the sky is dark should not be used. Instead, the images which have good lighting condition should be used for training.

Risk three: Crawling images from DSAT may have different sizes or types of images.

Solution: Images size and type are two critical parameters which should be unified before training them. Therefore, before training data, images from DSAT, should be classified in terms of place and unified in same size and type.

References

[1] *Transport Bureau of the Macao S.A.R.* (2018). Retrieved from Transport Bureau of the Macao S.A.R.: http://www.dsat.gov.mo

[2] Jia Wan, Y. Y. (2017). Traffic congestion analysis: a new perspective. IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 1398-1402). New Orleans, LA, USA: IEEE.

[3] Anıl Atvar, Y. O. (2017). Traffic density analysis using traffic camera images. *2017 25th Signal Processing and Communications Applications Conference (SIU)* (pp. 1-4). Antalya, Turkey: IEEE.

[4] *TensorFlow*. (2018). Retrieved from TensorFlow: https://www.tensorflow.org