**Refined Project Description**

The capabilities of humans executing complex tasks is rapidly being overtaken by artificial intelligence and machine learning. Conventional traffic light systems that are currently in place are inefficient and in many cases, generate more traffic congestion and they resolve. In my final year project, I will aim to address these challenges by comparing and contrasting various traffic congestion control systems that can be implemented using artificial intelligence and machine learning by focusing on increasing throughput, safety and efficiency. This is especially important because the number of autonomous vehicles is increasing faster year after year. By researching and constructing different AI approaches to this issue, I can answer the question of “Can Artificial Intelligence Reduce Traffic Congestion in Cities” myself, reinforced by my own research, statistics, implementations, simulations and conclusions.

Below are the provisional steps I will take to achieve this:

* Firstly research (or develop) various different AI based traffic congestion solution models
  + This will involve a phase of gathering a large amount of data - Dublin City Council
  + Develop my own based on data gathered?
* Compare and contrast them based on their throughput of roads, safety and efficiency to determine the most successful models
* Then I can compare the idealistic models to current outdated solutions that currently in place in a set of initially chosen cities e.g {Dublin, London, Paris}
* I will then develop a visual simulation in order to aid myself in analysing, refining and eventually showcasing my solutions
  + What software?
* It is vitally important to quantitatively evaluate my algorithms based on the criteria previously mentioned by simulating them
  + Scenario-Based Testing? → Rush Hour, Special Events, Emergencies
* Write about my findings and conclusions in my report
* Conference Paper?

Traffic congestion forecasting has two basic steps:

* Data collection, processing and cleaning:
  + Missing or null values
    - Dropping rows
    - Replacing null values with approximates (Median, Average, etc)
  + Detecting duplicate data
    - Potential dropping of rows to prevent noise
  + Detecting outliers
  + Replacing datatypes for efficiency
* Prediction model development:
  + Developing the right classifiers and models