

[< Back to Machine Learning Engineer Nanodegree](#)

Capstone Proposal

REVIEW

HISTORY

Meets Specifications

Dear student

This is the same reviewer from last time. Great job updating your proposal! I think that you're meeting the specifications and it's clear that you have a pretty good idea of what you want to do. Your suggestions are all feasible and I think you're on the right track.

About the capstone project:

While the code and implementation are both important, keep in mind that the capstone report is the most important element of your final project. This report simulates a formal submission to a journal for peer-review. Publishing your results is a key credential in machine learning and we want you to be ready for this!

You should have very little trouble quickly passing the project review if you initially follow the report template. Keep in mind that reviewers will be looking to see that you not only fully document how you implemented your project, but why you made the choices you made. This type of critical thinking is crucial to publishing in a peer-reviewed journal. Based on your proposal, I don't think you'll have much trouble with this, but I wanted to mention it up front.

I think you're definitely on solid ground and you've picked an interesting topic for your project. I'm looking forward to seeing the final result!

Cheers!

Project Proposal



Student briefly details background information of the domain from which the project is proposed. Historical information relevant to the project should be included. It should be clear how or why a problem in the domain can or should be solved. Related academic research should be appropriately

cited. A discussion of the student's personal motivation for investigating a particular problem in the domain is encouraged but not required.

Excellent! The resource you cited takes a great approach to using Cartesian coordinates to predict clusters of incidents. As noted before, you could definitely save some time and re-use this section as the **Project Overview** in your capstone report (just include a link to the dataset).



Student clearly describes the problem that is to be solved. The problem is well defined and has at least one relevant potential solution. Additionally, the problem is quantifiable, measurable, and replicable.

The problem would be a regression. Instead of predicting the number of emergencies expected overall, I have planned to predict 3 variables - the number of emergencies expected in each category - EMS, Fire and Traffic. The inputs would be the geographic location, date and time. The dataset does not explicitly include a column to indicate the count per emergency category. So, it will be rearranged accordingly.

Nice work! This is a very clear statement of the problem.



The dataset(s) and/or input(s) to be used in the project are thoroughly described. Information such as how the dataset or input is (was) obtained, and the characteristics of the dataset or input, should be included. It should be clear how the dataset(s) or input(s) will be used in the project and whether their use is appropriate given the context of the problem.



Student clearly describes a solution to the problem. The solution is applicable to the project domain and appropriate for the dataset(s) or input(s) given. Additionally, the solution is quantifiable, measurable, and replicable.



A benchmark model is provided that relates to the domain, problem statement, and intended solution. Ideally, the student's benchmark model provides context for existing methods or known information in the domain and problem given, which can then be objectively compared to the student's solution. The benchmark model is clearly defined and measurable.

A linear regression is a fine baseline. Sometimes a simple model like this will perform strongly, or sometimes it will suffer from bias and just not work well. Another good approach (if you're looking for a backup benchmark model) would be to simply use the number of calls from the previous day (or from the same day of the previous week/month) as the prediction.



Student proposes at least one evaluation metric that can be used to quantify the performance of both the benchmark model and the solution model presented. The evaluation metric(s) proposed are

appropriate given the context of the data, the problem statement, and the intended solution.



Student summarizes a theoretical workflow for approaching a solution given the problem. Discussion is made as to what strategies may be employed, what analysis of the data might be required, or which algorithms will be considered. The workflow and discussion provided align with the qualities of the project. Small visualizations, pseudocode, or diagrams are encouraged but not required.

Great job! I think you've clearly explained how you'll handle the data for pre-processing and splitting.



Proposal follows a well-organized structure and would be readily understood by its intended audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used and referenced are properly cited.

Everything is still well written and clear.

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)