



PROJECT

Capstone Proposal

A part of the Machine Learning Engineer Nanodegree Program

PROJECT REVIEW

CODE REVIEW

NOTES

SHARE YOUR ACCOMPLISHMENT!  

Meets Specifications


Great job with the proposal! I'm excited for you as you proceed with your work on completing the capstone.

Keep in mind that reviewers of your capstone will be checking that you not only document **how** you implemented your code but also **why** you did so. If you follow the [report template](#), you should be able to quickly pass the final review.

Best of luck to you with the project! 

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.

Good work outlining the project and providing background information on the online sentiment problem domain. This is definitely an interesting problem that I haven't seen be tackled by an MLND student before. Good choice! 

To complete this section for the final report, just be sure to also include references/links to any existing research in this field.

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.

Nice job defining the classification problem, and discussing how the solution needs to perform the sentiment analysis.

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.

Good discussion of the movie reviews data, including some helpful summary stats and descriptions of what the dataset contains.

Just make sure to include some samples of the data directly in your final writeup. This is an important part of the [exploratory analysis](#) and helps readers understand just how tricky a problem this is.

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.

Nice job proposing a general outline of a solution for the problem and identifying how the trained model will also analyze the twitter data. 😎

You may have already looked into these, but here are some [NLP](#) resources that might be worth exploring:

- <https://elitedatascience.com/python-nlp-libraries>
- <https://explosion.ai/blog/deep-learning-formula-nlp>
- <http://ruder.io/deep-learning-nlp-best-practices/index.html>

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.

Great work providing a reasonable benchmark for the project using the published result, although the final report should also describe what kind of model produced this "baseline" score. 😊

As an alternative, you could also fit a simple [linear classifier](#) to demonstrate how a final solution needs to introduce non-linearity to produce a good result.

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.

Good work describing the metrics you'll use to evaluate the model's performance, and showing the diagram of how the confusion matrix values are calculated.

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 discussion of a workflow for solving the sentiment classification problem. Let's hope you find it challenging and rewarding as you complete the implementation. 😊

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.

 [DOWNLOAD PROJECT](#)

[RETURN TO PATH](#)

Rate this review

[Student FAQ](#)