UDACITY

Capstone Proposal

REVIEW
HISTORY

Meets Specifications

Great job with the proposal! I'm excited for you as you start making predictions and complete your work on 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 bank marketing problem domain.

This is an interesting real world business problem that can definitely be tackled with machine learning. 😄

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 predict the clients' subscription behavior.

To expand here for the final writeup, you can also briefly mention your overall strategy for solving the problem as well, along with what algorithms/approaches were used. 😉

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 data from UCI, including some helpful descriptions of what the dataset contains.

Just make sure your final report also provides important summary stats (e.g., feature means/std devs, etc) and try to include some samples of the data directly in your 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 and identifying supervised learning algorithms and grid search techniques to be used. 😎

If you haven't considered them already, xgboost and LightGBM are popular gradient boosting models that could be appropriate to use here.

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 one of the untuned models.

To simplify the benchmark, you could also use a naive predictor that just predicts the majority class — this is a good sanity check for deciding whether your additional learned models are producing any meaningful output.

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 accuracy metric you'll use to evaluate the model's performance, and showing the equation used to calculate the scores.

If the target classes happen to be imbalanced, it would be best to also use an F-score or AUC/ROC metric as well. 😉

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 binary classification problem. I hope you find it challenging and rewarding as you complete the implementation. 😃

Another idea would be to try combining multiple models together with something like a VotingClassifier, or look at stacking models together to improve the prediction performance.
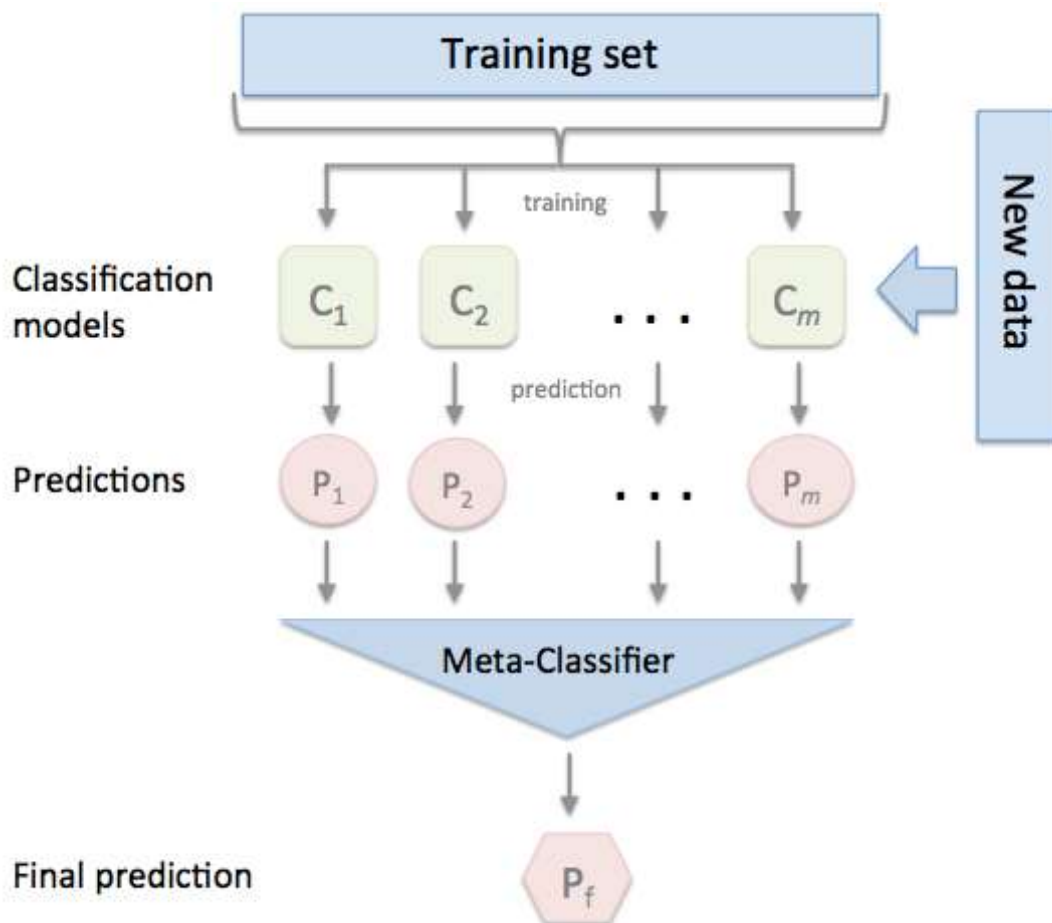
Proposal follows a well-organized structure and would be readily understood by its intended audience. Each section is written in a clear, concise and specific manner. Few grammatical and spelling mistakes are present. All resources used and referenced are properly cited.

↓ DOWNLOAD PROJECT

RETURN TO PATH

Rate this review

Student FAQ