# Homework Assignments
# Presentation/Report Guidelines

Course: Machine Learning in Marketing

Current as of 2020-12-09

## Presentation Outline

### Introduction

Start with a brief summary of the problem you are trying to solve. This ensures that your audience has the necessary context to understand (1) your approach and (2) the results you present. You can also motivate why the problem is relevant.

### Problem Formalization

A key part of your presentation (and your project) is translating the business problem into a machine learning task. This includes a mathematical formalization of the learning problem. First, you need to define the target variable. Do you tackle the problem by supervised or unsupervised learning? Do you need to post-process model predictions to derive your final results? This step is not only the basis for applying your models to solve the business problem but it also guides many of the subsequent decisions (e.g., selecting the right model for your task).

Second, you need to select the learning algorithm(s). This decision is directly based on how you define the target variable. Do you plan to have multiple sequential model stages? Are multiple models required to derive your target variable? Do you plan to stack models? Think about what approaches are known to work well in your context ("existing literature").

Third, define how you construct the training and validation set. What is the nature of your prediction task: Is it a time-series problem or do you predict cross-sectional information? How can you evaluate your model to ensure that it generalizes well to unseen data? How will you tune hyperparameters or train stacked models?

Fourth, explain what features you use in your implementation. Go beyond explaining how you build (or "engineer") the features and explain why you think that the proposed features are necessary. Can you link features to existing theory? In your opinion, what information will the features capture? Can you back your claims/expectations with EDA? A similar reasoning holds for architectural decisions in models (e.g., neural networks). How can you encode data patterns in your model's architecture?

### Description of baselines

Briefly describe the baselines you are using to provide context to your findings. How do you implement the baselines? Why are the proposed baselines relevant? For example, do you use baselines to verify that your pipeline works as expected or are baselines the current state-of-the-art in industry or research? What insights can you derive from the baselines' results?

### Results

Present the key results and supplementary findings that confirm your key results. Are the results in line with your expectations? Can you identify areas where your approach does especially well or where your model fails? Can you explain which features/architectural choices contribute to good performance? What data is most useful? Add a discussion of your findings.

### Conclusion and next steps

Summarize your approach and the key results. Can you make recommendations for marketing applications? Highlight limitations of your approach and provide ideas for next steps (future research).

## Presentation Format

You are free to choose any (suitable) presentation format when presenting your homework assignments. You can present the results using slides, a jupyter notebook, or in another suitable format. Regardless of the format, follow the structure outlined above.

Please keep in mind that we have scheduled 7-10 minutes for each presentation, followed by 2-3 minutes Q&A.

## Report Format

The report for the final class project should summarize your approach and results on 7-10 pages (including figures and tables). You can provide additional necessary information in a short appendix. The report's content should follow the structure outlined above.