**Course Six**

# The Nuts and Bolts of Machine Learning



# Instructions

Use this PACE strategy document to record decisions and reflections as you work through the end-of-course project. As a reminder, this document is a resource that you can reference in the future and a guide to help consider responses and reflections posed at various points throughout projects.

# Course Project Recap

Regardless of which track you have chosen to complete, your goals for this project are:

* Complete the questions in the Course 6 PACE strategy document
* Answer the questions in the Jupyter notebook project file
* Build a machine learning model
* Create an executive summary for team members and other stakeholders

# Relevant Interview Questions

Completing the end-of-course project will empower you to respond to the following interview topics:

* What kinds of business problems would be best addressed by supervised learning models?
* What requirements are needed to create effective supervised learning models?
* What does machine learning mean to you?
* How would you explain what machine learning algorithms do to a teammate who is new to the concept?
* How does gradient boosting work?

**Reference Guide:**

This project has seven tasks; the visual below identifies how the stages of PACE are incorporated across those tasks.



**Data Project Questions & Considerations**

**PACE: Plan Stage**

* What are you trying to solve or accomplish?

We are trying to increase revenue for the taxi company by predicting customer behavior.

* Who are your external stakeholders that I will be presenting for this project?
* What resources do you find yourself using as you complete this stage?

Besides the documentation, guidelines and workflows in the course itself. I am using search engines and AI chatbots to further understand wherever I get stuck.

* Do you have any ethical considerations at this stage?

I am deeply concerned that informing the driver if the customer is likely to tip or not, specially if it is a false positive, will influence driver’s behavior toward the customer. Their attitude may become reckless which in effect will impact the company’s reputation.  
In case of false positive, a customer might not tip if he is not pleased with the attitude of the driver.

* Is my data reliable?

I suppose yes. At this point I must take the data at it’s face value and assume it is coming from a reliable team.

* What data do I need/would like to see in a perfect world to answer this question?
* What data do I have/can I get?
* What metric should I use to evaluate success of my business/organizational objective? Why?

**PACE: Analyze Stage**

* Revisit “What am I trying to solve?”Does it still work? Does the plan need revising?
* Does the data break the assumptions of the model? Is that ok, or unacceptable?
* Why did you select the X variables you did?
* What are some purposes of EDA before constructing a model?
* What has the EDA told you?
* What resources do you find yourself using as you complete this stage?

**PACE: Construct Stage**

* Do I notice anything odd? Is it a problem? Can it be fixed? If so, how?
* Which independent variables did you choose for the model, and why?
* How well does your model fit the data? What is my model’s validation score?
* Can you improve it? Is there anything you would change about the model?
* What resources do you find yourself using as you complete this stage?

**PACE: Execute Stage**

* What key insights emerged from your model(s)? Can you explain my model?
* What are the criteria for model selection?
* Does my model make sense? Are my final results acceptable?
* Do you think your model could be improved? Why or why not? How?
* Were there any features that were not important at all? What if you take them out?
* What business/organizational recommendations do you propose based on the models built?
* Given what you know about the data and the models you were using, what other questions could you address for the team?
* What resources do you find yourself using as you complete this stage?
* Is my model ethical?
* When my model makes a mistake, what is happening? How does that translate to my use case?