**Capstone Project 1: Proposal**

This first capstone project will involve creating a model to score leads. Specifically we want to understand how valuable a lead is based on certain characteristics about that lead.

My client for the project is my current company, a B2B Saas company interested in scaling sales operations. Currently we use a third-party vendor for modeling and lead scoring. A significant amount of effort and financial investment has gone into the third-party tool and it would be great if we could benchmark the value that tool provides.

For this project we’ll be using lead and opportunity data from Salesforce, stored in an AWS Redshift warehouse.

Lead scoring is a classic candidate for machine learning. We want to provide a numeric score that indicates the quality of the lead based on meta-data about the lead (person, lead source, campaign, associated company, etc). We have 4 years of consistent data spanning thousands of accounts and interactions with customers and prospects.

The resulting deliverables are:

* An initial model
* Code
* A write-up detailing the requirements, methodology and resulting model
* Presentation to my manager

Resources & Research:

* <https://www.blendo.co/blog/lead-scoring-data-science/>
* <https://github.com/turi-code/tutorials/blob/master/dss-2016/lead_scoring/lead_scoring_tutorial.ipynb>
* <https://www.periscopedata.com/blog/simple-lead-scoring-with-enrichment> ⇒ not python but interesting how they conducted the analysis
* <https://www.analyticsvidhya.com/blog/2015/09/build-predictive-model-10-minutes-python/>
* Business context around lead scoring: <https://labs.openviewpartners.com/innovative-lead-scoring-models/#.W8aAQUtKiUk>
* General process of building lead scoring system: <https://blog.hubspot.com/marketing/lead-scoring-instructions>
* Algorithms: <https://ai.stackexchange.com/questions/4601/which-algorithm-should-i-choose-for-lead-scoring>
* Example Repo: https://github.com/xeneta/LeadQualifier
  + Using NLP to qualify: <https://medium.com/xeneta/boosting-sales-with-machine-learning-fbcf2e618be3>
* How insight utilized into business: <https://www.urbanairship.com/blog/churn-prediction-our-machine-learning-model>
* <https://hackernoon.com/supervised-machine-learning-linear-regression-in-python-541a5d8141ce>
* <https://dataconomy.com/2015/05/predictive-machine-learning-behind-the-scenes-at-fliptop-and-predictions-for-the-future-of-martech/>
* <https://machinelearningmastery.com/spot-check-machine-learning-algorithms-in-python/>

Used in order to define:

* Machine learning model
* Variables
* Target value
* Data Structure (i.e. rows, etc)

### **Project: Capstone Project 1: Project Proposal**

Finalize one capstone idea based on the feedback that you received on your initial ideas and your discussions with your mentor.

**Submission instructions**:

1. Write your proposal in a Google Doc (1-2 pages) and submit the link via the “Submit” button. Make sure your mentor has permissions to comment on the document.
2. Work with your mentor to incorporate any feedback into later drafts and submit as many times as needed.
3. Once your mentor has approved your proposal, convert the doc to a PDF file.
4. Create a GitHub repository for this project (if you haven’t done so already).
5. Add the PDF to your GitHub repository for this project.
6. Share the proposal with your peer community for feedback.

Note: All code and further documentation you write will be added to this repository.