Submission

Project Title \*

​

Which topic did you choose to apply the data science methodology to? **(2 points)**

Enter text here﻿

Emails

Next, you will play the role of the client and the data scientist.

Using the topic that you selected, complete the Business Understanding stage by coming up with a problem that you would like to solve and phrasing it in the form of a question that you will use data to answer. **(3 points)**

**I work with logistics and I have an absurd amount of emails coming in and out. Generally, they are treated by the customer service. However, there's one special case where no mistake can be made, which are WPS (wrong product shipped) this is configured by shipping the wrong product, shipping the right product to the wrong destination, shipping product on a different date than required by customer or any other case where there's a difference between customer's plan and product shipping. As this is a constant KPI, we have actual data on WPS and “Near misses” where the email came, was overlooked, but human intervention didn’t let it turn into a WPS. Is it possible to classify incoming emails by its potential of ending in a WPS and warn Customer Service so they pay extra attention and avoid mistakes happening due to overllooking?**

You are required to:

1. Describe the problem, related to the topic you selected.
2. Phrase the problem as a question to be answered using data.

For example, using the food recipes use case discussed in the labs, the question that we defined was, "Can we automatically determine the cuisine of a given dish based on its ingredients?".

Enter text here﻿

Your answer needs to be a little bit longer. Write a few sentences to complete your assignment.

Briefly explain how you would complete each of the following stages for the problem that you described in the Business Understanding stage, so that you are ultimately able to answer the question that you came up with. **(5 points)**:

1. Analytic Approach

**The goal is to classify emails. This could take two approaches:**

**Classification: Yes/No for potential WPS**

**Predictive: Classify the level of threat. Choose a scale and classify the email depending on the chance of turning into a WPS**

1. Data Requirements
2. Data Collection
3. Data Understanding and Preparation
4. Modeling and Evaluation

You can always refer to the labs as a reference with describing how you would complete each stage for your problem.

Enter text here﻿