Springboard Data Science Career Track
Capstone Project 1
Charles Tremblay

Project Proposal

1. Problem

Superconductors offer extraordinary electrical conductivity with zero resistance. They are very useful in applications that require strong magnetic fields or long distance power transmission.The benefits of superconductors are clear, but are restricted by their very low critical temperature as it is far too low to be used in a real world setting. In addition, superconductivity is not a well understood phenomenon, which makes it difficult to create a superconductor with a desired critical temperature. To aid researchers study superconductivity, a model can be developed from material properties of existing superconductors with the goal of predicting its critical temperature. This can help in the study and development of new superconductors by offering insight on what properties of the material can influence its critical temperature.

2. Client

The prospective clients would be researchers who want to develop new superconducting material and better understand superconductivity. The developed model can help them analyze their data and focus their time on more promising materials. Companies developing technology that utilizes superconductors could also benefit from this model. For example, MRI machines, particle accelerators, and digital circuits. The benefit of this model could help companies develop technologies in a more affordable and scalable fashion by choosing the appropriate material.

3. Dataset

The dataset was retrieved from the UCI machine learning repository. The data set contains two files; one detailing the chemical composition of each unique superconductor, the other containing the material properties and critical temperature of the superconductor. Overall there are 21,264 superconductors with 82 attributes in the dataset. Link: http://archive.ics.uci.edu/ml/datasets/Superconductivty+Data

4. Steps to Solve

The first step will be to perform data wrangling and EDA for exploration of data. Then I will apply statistical methods for hypothesis testing. Finally. I'll use machine learning to develop models that will be able to predict the critical temperature of a given superconductor.

5. Deliverables:

Code, report, and a slide deck for the Github repository.