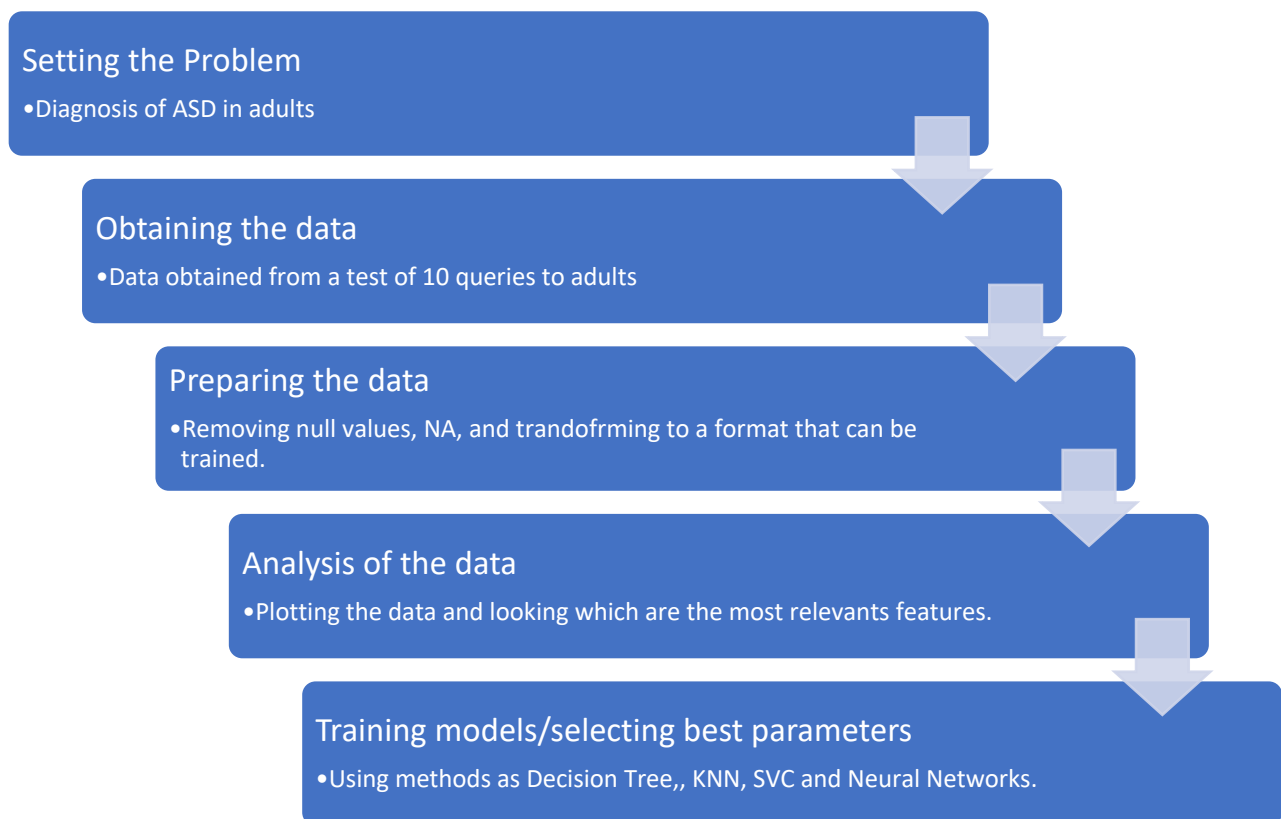


Raul Barrantes Pampillo  
Cenfotec  
Learned Lessons Report  
Data Science with Python  
Task 4: Complete a Data Science Capstone Project  
Capstone Report

## Problem to solve

Detecting and diagnosing the Autistic Spectrum Disorder can be lengthy and difficult, especially in adults. Autistic Spectrum Disorder (ASD) is a neurodevelopment condition associated with significant healthcare costs, and early diagnosis can significantly reduce these. A dataset is provided with 20 features: 10 as answers, and the other 10 with attributes related to the description of the patients.

## Data science process



## Best practices

- Prepare the data/ it is preferable to have null data as NAs.
- Converting categorical features using encoding methods.
- Testing different methods, models with respective values and hyperparameters.
- Remove values that can cause overfitting.

## Analysis of methods used

	Hyperparameters	Hyperparameters selected	Accuracy – Testing data-
Neural Networks	<b>min_samples_split:</b> 2-10	<b>min_samples_split:</b> 3	93.61%
SVC	<b>n_neighbors:</b> 1-4 <b>weights:</b> uniform, distance'	<b>n_neighbors:</b> 3 <b>weights:</b> uniform	95.39%
KNN	<b>C:</b> 0.5, 2, 0.5	<b>C:</b> 1.5	95.74%
Decision Tree	<b>hidden_layer_sizes:</b> 1-10 <b>alpha:</b> 0.0001, 0.001, 0.01	<b>hidden_layer_sizes:</b> 1-10 <b>alpha:</b> 0.001	98.58%

## Main Lessons Learned

- Using a tool like Python with Scikit-Learn can provide us a greater insight into Data Analysis.
- In addition to this, the use of different types of a plot can provide an explanation for the different data.
- Also, it is important to transform and clean the data, because most of it at first glance it is not ready to perform any kind of analysis/exploration/prediction.
- It is important to test and select the best hyperparameters for each method.