What is problem-solving?

While there are many definitions of "*problem-solving*" perhaps the most applicable one is:

Problem-solving is what you do when you don't know what to do.

Now, this may seem silly. But almost all other definitions are circular in some way. They basically say something like: "*problem-solving is the act of solving problems.*" They might use fancy language in doing this, but they are generally circular!

As an exercise, see if you can find definitions to "*problem-solving*". You can use any source, AI, websites, books, articles, etc. Can you see the circular nature of the definitions?

A problem-solving framework

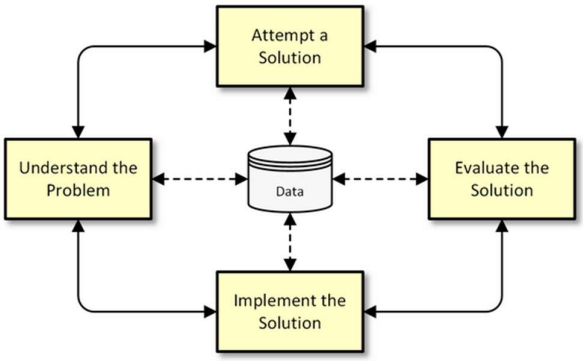
Here we will give some guidelines on the process of problem-solving, in the broad context of data science, machine learning and artificial intelligence applications. It is important to note that problem-solving often requires:

- a systematic process of how to approach problems,
- an application of a broad spectrum of skills, tools and techniques,
- an iterative process,
- critical thinking capabilities, and
- perseverance and "grit".

A process for solving a data-science, AI or machine-learning problem consists of iteratively:

- Trying to ***understand*** the problem
- ***Attempting*** a solution to the problem
- ***Evaluating*** the solution
- ***Implementing*** the solution

This is represented by the following diagram:



CRISP-DM as a framework for problem-solving

The problem-solving process outlined above, provides a conceptual basis for the CRISP-DM framework you have already seen.

Understanding the problem

The first, and most important step in problem-solving is understanding the problem. In the context of data science and machine learning, this involves:

1. ***Formulating the Business Problem:*** This is a crucial first step. It is necessary to clearly define the problem one is trying to solve. What is the question you trying to answer? What decision needs to be made? What assumptions are there? This focuses the approach and usually helps in selecting appropriate tools and techniques.

2. ***Understanding the Data:*** Having formulated the problem, one needs to identify the appropriate data that can lead to a potential solution. This might involve collecting new data or using existing datasets. Understanding the data's source, structure, quality, format, and any limitations is essential.

Attempting a solution

When attempting a solution an important step is selecting the appropriate machine learning algorithm for the problem. This requires:

3. ***Preparing the Data:*** Data needs to be processed, manipulated and cleaned for the appropriate model being used to solve the problem.

4. ***Model selection:*** One needs to select and implement an appropriate mathematical or machine-learning model as appropriate.

Evaluating the solution

After having made an attempt at solving the problem, one needs to determine if the solution actually solves the problem one is trying to solve. This is the step 5 in the CRISP-DM workflow,

- *Does it provide a suitable answer to the initial the question you trying to answer?*
- *How accurate is the solution?*
- *Are there possible alternatives approaches?*
- *Does it address the decisions that needs to be made?*
- *Are the initial assumptions still applicable?*
- *How do different models compare to each other?*

In evaluating the solution(s) one should assess the applicability, feasibility, pros and cons of the possible solutions. This process should aid in selecting the most appropriate solution in cases where multiple solutions are available,

Implementing the solution

If the model meets the success criteria (i.e. solved the problem), it can be deployed for real-world use. This is the step 6, ***Deployment***, in the CRISP-DM workflow, and might involve integrating it with an existing application or service. One needs to consider aspects like scalability, monitoring, and potential biases in the model.

Often one needs to modify the solution to create a pipeline for deployment. This aspect will not be covered in this course.