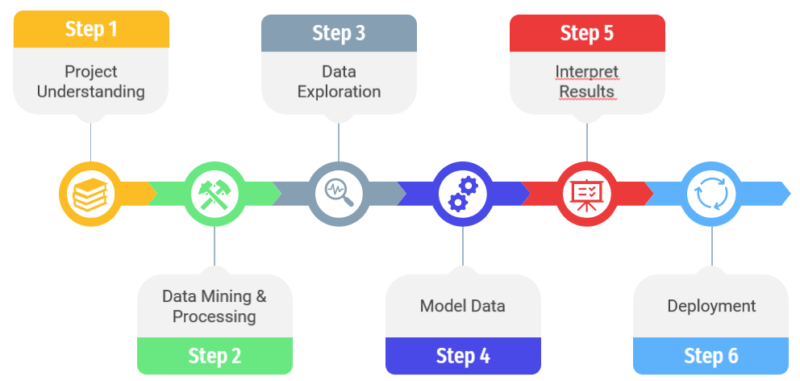
Proposal: Metis Data Science Project

1. Introduction

Understanding the business or activity that your data project is part of is key to ensuring its success and the first phase of any sound data analytics project. To motivate the different actor’s necessary to getting your project from design to production; your project must be the answer to a clear organizational need. Before you even think about the data, go out and talk to the people in your organization whose processes or whose business you [aim to improve with data](https://blog.dataiku.com/how-will-machine-learning-and-ai-change-my-organization). Then, sit down to define a timeline and concrete key performance indicators. I know, planning and processes seem boring, but, in the end, they are an essential first step to kick start your data initiative!

* Data acquisition, cleaning, and aggregation
* Exploratory data analysis and visualization
* Feature engineering
* Model creation and validation
* Basic statistical and mathematical foundations for data science



Therefore in this proposal we know how to used Data Analysis and apply Machine Learning Algorithm’s. So the project is or problem is sub class or sub models so over project is pieces

That way we start.

1. **Question/need:**

Spread denial of service attacks pose a significant threat to network security and cause significant disruption in the network with malicious traffic. Although many statistical methods for detecting distributed denial of service attacks have been developed, designing a real-time detector with low computational overhead remains one of the primary concerns. The evaluation of new detection algorithms and techniques, on the other hand, is heavily reliant on the availability of well-designed datasets. It can exploit attacks in a variety of ways, as illustrated by the image below, which depicts a bright idea of how these attacks occur:

1. What is the framing question of your analysis, or the purpose of the model/system you plan to build?
2. Who benefits from exploring this question or building this model/system?
3. **PROJECT COMPONENT**

**3.1 Design:**

This problem a design is Laptop price now days in markets it is real world problem

It’s depended lots of features Define Project Goal. ...

1. Determine Outcomes, Objectives, and/or Deliverables. ...
2. Identify Risks, Constraints, and Assumptions. ...
3. Prepare a Visual Aid. ...
4. Ballpark Your Budget. ...
5. Determine Approval and Monitoring Processes. ...
6. Use Proper Project Design Documents.

**3.2 Dataset:**

If you’re working on a personal project or playing around with a dataset or an API, this step may seem irrelevant. It’s not. Simply downloading a cool open dataset is not enough. In order to have motivation, direction, and purpose, you have to identify a clear objective of what you want to do with data: a concrete question to answer, a product to build,

**Kaggle and UCL**

Which contain a lots of data one is Laptop price prediction at csv file format

Once you’ve gotten your goal figured out, it’s time to start looking for your data, the second phase of a data analytics project. Mixing and merging data from as many data sources as possible is what makes a data project great, so look as far as possible.

**Clean Data:**

The next step (and by far the most dreaded one) is cleaning your data. You’ve probably noticed that even though you have a country feature, for instance, you’ve got different spellings, or even missing data. It’s time to look at every one of your columns to make sure your data is homogeneous and clean.

Now that you have clean data, it’s time to manipulate it in order to get the most value out of it. You should start the data enrichment phase of the project by joining all your different sources and group logs to narrow your data down to the essential features. One example of that is to enrich your data by creating time-based features, such as:

* Extracting date components (month, hour, day of the week, week of the year, etc.)
* Calculating differences between date columns
* Flagging national holidays

1. **Algorithms:**

When collecting, preparing, and manipulating your data, you need to be extra careful not to insert unintended bias or other undesirable patterns into it. Indeed, the data that is used in building machine learning models and AI algorithms is often a representation of the outside world, and thus can be deeply biased against certain groups and individuals. [One of the things that make people fear data and AI](https://blog.dataiku.com/finding-hope-in-artificial-intelligence) the most is that the algorithm isn’t able to recognize bias. As a result, when you train your model on biased data, it will interpret recurring bias as a decision to reproduce and not something to correct.

* [1. Classification Algorithms](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#1_Classification_Algorithms)
  + [a) Naive Bayes](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_a_Naive_Bayes)
  + [b) Decision Tree](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_b_Decision_Tree)
  + [c) Random Forest](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_c_Random_Forest)
  + [d) Support Vector Machines](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_d_Support_Vector_Machines)
  + [e) K Nearest Neighbours](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_e_K_Nearest_Neighbours)
* [2. Regression Algorithms](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#2_Regression_Algorithms)
  + [a) Linear regression](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_a_Linear_regression)
  + [b) Lasso Regression](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_b_Lasso_Regression)
  + [c) Logistic Regression](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_c_Logistic_Regression)
  + [d) Multivariate Regression](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_d_Multivariate_Regression)
  + [e) Multiple Regression Algorithm](https://www.upgrad.com/blog/types-of-artificial-intelligence-algorithms/#_e_Multiple_Regression_Algorithm)

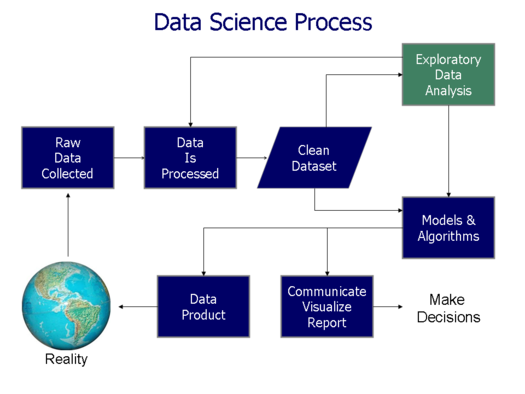
1. Tools used:

Jupiter notebook, Python 3.6 or 3.9

Numpy, pandas, seaborn, sklearn, Mathplotlib, a lot of more etc.

Annocda

Language: Python



1. **Outcomes**

* An understanding of problems solvable with data science and an ability to attack them from a statistical perspective.
* The ability to create data analytical pipelines and applications in Python.
* Familiarity with the Python data science ecosystem and the various tools needed to continue developing as a data scientist.

1. PPT Slide:

The summary of project point to point every phase written from phase to phase

