Applied Data Science Project

L5 - 10 practical tips

Giuseppe Rizzo

Turin, October 4, 2022

Customer centricity

- Does the project bring value to a customer?
    - Buy/use depends on the value

- Will the customer easily use the results of the project?
    - The first customer is "who develops the project"

- Can the project be executed?
    - Start from a PoC and assess feasibility (technically, financially)

- Does it make sense for the business of the venture?
    - Alignment with venture strategy

Customer centricity

- Does the project bring value to a customer?
  - Buy/use depends on the value

- Will the customer sults of the project
  - The first customer elops the project

- Can the project
  - Start from a PoC and assess feasibility (technical)

- Does it make sense for the business of the venture?
  - Alignment with venture strategy

Customer feedback

Engage customers at regular intervals to gather their feedback

Think to a feedback in a constructive manner 100% of times even when it may seem out of scope or hard to tolerate

Focus on the outcome (i.e. what is the ultimate goal of the work) rather than the output (i.e. what has been done)

Customer feedback

Engage customers at regular intervals to gather their feedback

Think to a feedba seem out of scope

Focus on the out output (i.e. what it has been done)

Fall in love with the problem and not the solution

Tip #3

Solve creatively the problem

Be eager to iterate back from requirements to implementation and backward

Keep it simply, not everything needs to be solved with a neural net, most of the times a random forest makes the trick

Focus on data and historical samples

Fall in love with the problem and not the solution

The company was born to develop devices to entertain people

Initially with cards than with consoles

Be aware of risks and monitor achievements

Tip #4

- Define what are the risks to conduct the project

- Define strategies to mitigate the risks

- Monitor achievements

- Take time to recognize the achievements and to share them with your customers this is fuel for addressing the forthcoming challenges while reassure your clients that the project is real

Share the burden

Projects turn to be successful if they are shared

Share responsibilities with an agreement based on analytical and objective evidences (who's more expert in development, who's in reporting, who's in design)
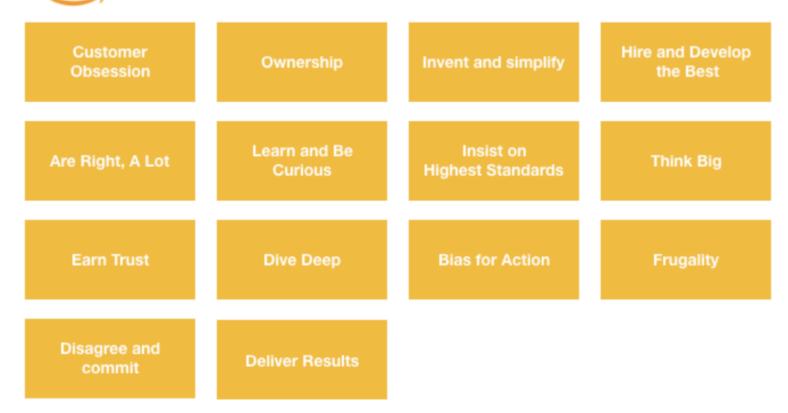
Anybody in the team with clear and leadership roles

Share the burden

amazon 14 leadership principles

Customer Obsession	Ownership	Invent and simplify	Hire and Develop the Best
Are Right, A Lot	Learn and Be Curious	Insist on Highest Standards	Think Big
Earn Trust	Dive Deep	Bias for Action	Frugality
Disagree and commit	Deliver Results		

Leaders care of the projects for any detail, which turns out to be successful for realizing an impactful output

Data Science is interdisciplinary

Tip #6

Activities are diverse

Better to have a multidisciplinary team (people covering different skills)

Work interdisciplinary

Diversity in both background and skills is key

Represent the project with a diagram

Represent the whole value chain from needs, data acquisition, what is done, till customers with blocks

Have self-explanatory blocks that each has an input arrow and an output arrow. Arrows are meant to define the requirements for the execution and the expected output

Be clear which block needs to interact with external APIs or internal pipes

Projects move data (in/out) thus reason in terms of data in/data out rather than source code in/source code out

A clear and well defined diagram speeds up the development as well as eases the communication with clients

Start with some code, visualize it, do folding

Tip #8

Code is an obscure matter for many, even for developers

Start iteratively with a visual tool for coding bits and visualize intermediate results

Though "black on white" is fascinating while programming, it turns out to be a barrier for communication to team members

Notebooks help in having a visualization that encompasses both coding and intermediate results

When the project becomes stable and the team is large and each has dedicated tasks then you can be back on "black on white"

Start with some code, visualize it, do folding

Tip #8

Avoiding having spaghetti coding

Structure coding in sections and describe with one sentence each section
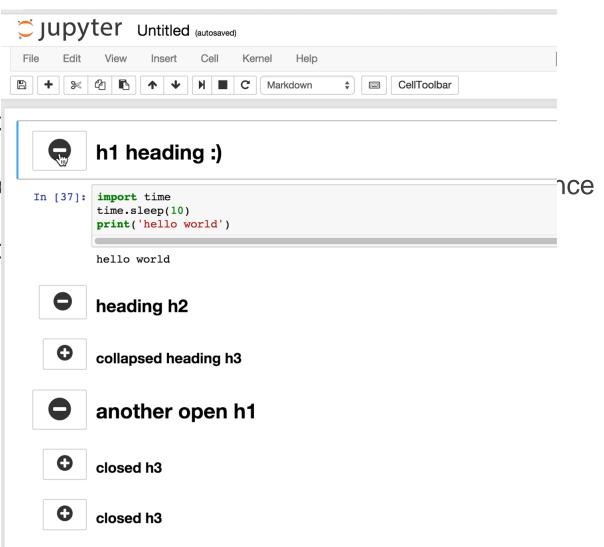
Do code folding in notebooks

Start with some code, visualize it, do folding

Tip #8

Avoiding having spaghett...

Structure coding in sectio... ...nce

Do code folding in noteb...

linksfoundation.com
COPYRIGHT ©2021 LINKS

Models keep running

Losing sight on what is going on in the project

Add notification strategies at certain events such as:
- the model is generated
- the model has been tested
- the model has encountered an issue

Models keep running

Losing sigh

Add notific
- the mo
- the mo
- the mo

Knock Knock

Made with Python | downloads 69k | downloads/month 1k | Star 2.2k

A small library to get a notification when your training is complete or when it crashes during the process with two additional lines of code.

When training deep learning models, it is common to use early stopping. Apart from a rough estimate, it is difficult to predict when the training will finish. Thus, it can be interesting to set up automatic notifications for your training. It is also interesting to be notified when your training crashes in the middle of the process for unexpected reasons.

Slack

Similarly, you can also use Slack to get notifications. You'll have to get your Slack room webhook URL and optionally your user id (if you want to tag yourself or someone else).

Python

```python
from knockknock import slack_sender

webhook_url = "<webhook_url_to_your_slack_room>"
@slack_sender(webhook_url=webhook_url, channel="<your_favorite_slack_channel>")
def train_your_nicest_model(your_nicest_parameters):
    import time
    time.sleep(10000)
    return {'loss': 0.9} # Optional return value
```

You can also specify an optional argument to tag specific people: user_mentions=[<your_slack_id>, <grandma's_slack_id>].

Versioning and milestones

Tip #10

A project is a living creature, it is necessary to have regular "pictures" of the developments

Avoid versioning only when things are "marvellous" because they will never be

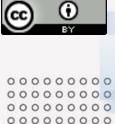Group versions into milestones to record the work done

Thank you for your attention.

Questions?

CONTACTS

Giuseppe Rizzo

Program Manager (LINKS Foundation) and
Adjunct Professor (Politecnico di Torino)

giuseppe.rizzo@polito.it