**Datascience activities roadmap proposal.**

**Introduction:**

Data insights are the beacon for decision makers to take the right measurement based on genuine and accurate data, at a certain proportion. Building, training and deploying models are the three main pillar for a ML (aka machine learning) project. Nonetheless, data wrangling is the first main step to span over before any data processing. By selecting and training the right competencies around; we are planning, by this project, to pave the path for a new Data scientist team to lead more important and audacious projects.

**National Exams Data insight:**

The project is intended to process Students’ data and get insight out of their achievement during both school year and national Exam.

As information does not come structure all the time, Data cleaning takes more than 80% of total planning project schedule. And here comes the power of Cloud services that may shorten the project implementation lifespan for data cleansing, training and even deploying.

National Exam data project will deliver different comparison charts to scrutinize tiniest variations and explore their impact of the whole assessment process.

Another level of processing will deliver a Prediction and classification for students’ achievement to have an accurate feedback on the future trend regarding a such endeavor.

**National Exam Project schedule (in days) :**

|  |  |  |  |
| --- | --- | --- | --- |
| Data Query&Graph | Data cleansing | Training /modelling | Test& deployment |
| 4 | 8 | 4 | 4 |

**Actual Stage**: *Data cleaning*.

**Training:**

A series of workshops will be held I premises for those who are willing to enroll in Datascience activities. A schedule will be established within the second week of January 2020.

***Samir Amri***