ISSU0053 Data Science and Big Data Analytics **Report ID: 1043342**

UCL International Summer School for Undergraduates 2019

**Assessment I: Computer Practical Work and Write-up (50%)**

|  |  |  |
| --- | --- | --- |
| **Task** | **Mark** | Comments |
| T1: Explore the dataset | A |  |
| T2: Linear **/** logistic regression **/** LDA | B |  |
| T3: Decision Tree methods | B |  |
| **Overall task:** |  |  |
| Use of Figures | A |  |
| Introduction, Commentary, Summary | B |  |
| Coding Technique | A |  |
| Writing Standard | A |  |
| Structure & Presentation | B-C |  |
| **Overall:** | A-B  (65%) |  |

**Strengths:**

**Focus on improvement:**

- Introduction not clear – what are you trying to predict? I only understood it when I reached the conclusion section

- you write in your report: Catch\_Rate. Numerical variable indicating how easy is to catch a Pokémon when trying to capture it to make it part of your team. It is bounded between 3 and 255. The number of different values it takes is not too high notwithstanding, we can consider it is a continuous variable.

→ if its a continuous variable – why do you use NB and KNN which are for classification????

this makes me assume that you did not understand some of the main concepts covered in class and do not know which algorithms to use for which sort of problem

- large parts of your reports is just printed output which doesnt add any value and makes the report messy and hard to read

- report is hard to read due to weird division between html and doc file

- report in general messy and very hard to follow (all of a sudden a lda model appears towards the end with 0 explanation?)