|  |
| --- |
| Date: March 12, 2018 |
| Time: 1500 – 1630hrs  |
| Location: SIS Meeting Room 4-1 |
| Attended By: Josh Ho, Friedemann Ang, Prof Kam, Ms. Meena |

**Instructor Meeting Minute 04**

|  |  |  |
| --- | --- | --- |
| **No.** | **Agenda** | **Action By:** |
| 1 | Presentation of findings from interim exploratory analysis* Presented on which features were significant in our exploratory study
* Prof Kam emphasises the importance of doing proper data preparation, from checking of missing values to deriving new features
* We should also be looking at predicting failure as the response variable
 | Both |
| 2 | Obtain direction for deliverables of project* Prof Kam reminded us of the importance of data exploration
* In terms of final deliverable, Prof Kam wanted us to focus on identifying features that have a significant relationship to our response variable. Depending on the method, data may need to be further transformed
* A method proposed by Prof Kam would be partitioning, but we could explore other methods such as boosted trees and naïve bayes as well
* We should focus on the training set of data, and split it into training, testing and validation sets when performing our modelling methods
 | Both |

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Tasks Allocated** | **By (date):** | **Action By:** |
| 1 | Uploading of meeting minutes | 14 Mar | Josh |
| 2 | Refine data preparation steps | 16 Mar | Josh |
| 3 | Deriving new variables from text data | 16 Mar | Friedemann |
| 4 | Explore different methods of defining predictor-response relationship | 18 Mar | Both |