

**MEETING MINUTES #8– SUPERVISOR**

**Date/Time
Venue
Meeting with
Attendees**

27th March 2017, 04:00pm
SIS MR 4.5
Prakash Sukhwal
Aayush Garg, Prekshaa Uppin, Akshita Dhandhania

**Agenda**

Update on our progress from previous week’s meeting with Prof Kam. Showed complete documentations and EDA presentation.

|  |  |  |
| --- | --- | --- |
| Point made by | Notes | Follow – up Action |
|  |  | **Task** | **Person(s) Responsible** | **Due Date** |
| Aayush, Akshita, Prekshaa | 1. We showed the improved EDA – our supervisor gave us some feedback:
2. Do not use bubble charts. It is not well accepted in the industry.
3. The interpretation of the graph needs to be very clear
4. Observe Hana Owens, she has high eigenvector but low betweenness. Try and look at such employees who show unusual patterns, and try to understand why this is happening.
5. Upload all the PPTs you’ve created for the EDA
6. Explained our attempt at regression and the issues we are having: Prakash’s comments –
7. The difference between 1 and 2 is not the same as the difference between 4 and 5. Do not take average of ratings, it is wrong to say. You should show the modal rating or show your median etc. There is no continuity in this rating. Try classification algorithm.
8. Try and clean the data in such a way that the decimal value will uniquely show the strength of a unique pair.
9. Do not treat the rating as a continuous variable
10. Judge the algorithm based on the response variable
11. Try out different models to understand which is better
 | Create EDA document | Aayush | 29th March 2017 |
| Try classification algorithm – Random forest | Prekshaa | 30th March 2017 |
| Try classification algorithm – decision trees | Akshita | 30th March 2017 |
| Try neural networks | Aayush, Akshita, Prekshaa | 3rd April 2017 |
| 1. We explained to him about the data cleaning we did, regarding edge weights. In the scenario that the two connected nodes rate each other highly varying scores, the current edge weights will be inaccurate. So we decided to convert the directed graph to an undirected graph and take an average of the two scores (to each other) as the edge weight.
 |  |  |  |
| Prakash | Try comparing 2 months of email data for every new employee (December vs January) - check how interaction changed for new employees over their joining time till now.  |  |  |  |

*Meeting minutes prepared by Prekshaa Uppin
Minutes has been vetted by Aayush Garg, Akshita Dhandhania*