4th Oct 2022

See Distribution

**MINUTES OF 01/2022 DATA ANALYTICS LAB (DAL) GROUP PROJECT MEETING HELD ON 04 OCT 2022**

**Present:**

|  |  |
| --- | --- |
| Shi Chee Liang | Team Member |
| Bryant Philippe Lee | Team Member |
| Chen Yiman | Team Member |

**In Attendance:**

|  |  |
| --- | --- |
| Nil |  |

**Absent with Apologies:**

|  |  |
| --- | --- |
| Nil |  |

|  | **ACTION BY** |  |
| --- | --- | --- |
| **OPENING** |  |  |
| 1. The meeting commenced at 1410 hrs. |  |  |
| **ITEM 1 – CONFIRMATION OF MINUTES** |  |  |
| 1. Follow Up item. There were no prior minutes or follow up items. |  |  |
| **Item 2 – Discussion and selection of topic for dal group project** |  |  |
| 1. This item sought to select a topic to proceed for the DAL group project. 2. Approach. Team Member, Shi suggested for the team to go through the topics one by one and for the members to share their opinions on the topics before deciding. He also reminded the members that Prof Kam’s suggestion was that topic 1 and 4 was suitable for everybody while topics 2 and 3 were more suited for people with some finance background. 3. Topic 1. Team Member, Shi opined that this topic seemed to be better documented in terms of the data set as well as expected outcomes. He shared that the this could make the topic easier to work on. Team Member, Chen queried if the outcome was to make recommendations on land slide prevention. Team Member, Shi and Team Member, Lee replied that we should only focus on the identification of land slides as prevention was a policy decision function and does not fall under the scope of data analysis. 4. Topic 2. Team Member, Lee noted that this topic seemed to focus on clustering. Team Member, Shi concurred and asked for his opinion on the difficulty level compared to topic 1. Team Member, Lee opined that the difficult level should be about the same. Team Member, Shi also queried if clustering was predictive since the requirement of the project was to build predictive models. Team Member, Chen replied that it should be and added that we should first cluster the data points before trying to identify factors that would allow us to build a predictive model. 5. Topic 3. Team Member, Lee commented that this topic might be one of the easier ones to work on. Team Member, Shi disagreed, citing that there seemed to be a lack of data for which to work on. He queried on how one might derive fraud simply from billing data; was it simply matching how much electricity was consumed against how much money was paid? Team Member, Chen suggested that the team first look at what the data covered before jumping to conclusions. 6. Topic 4. Team Member, Lee shared that this topic was very similar to DAL assignment 2. Team Member, Shi concurred and added that the scope was much wider as multiple crops needed to be identified. Team Member Chen suggested however that it could be easier to work on due to the similarities between this topic and assignment 2. 7. Selection. The team took some time to explore the data sets and topic before coming back again to decide on the chosen topic. Topic 1 was eliminated due to the assessed work required to research and understand the land slides and it’s contributing factors. Topic 3 was Eliminated due to the perceived ambiguity and lack of data in the datasets. Topic 4 was eliminated as it was deemed to be too big in terms of scope as well as requiring additional work in terms of coding in python. The team agreed on Topic 2 as the chosen topic for DAL group assignment. |  |  |
| **Item 3 – Projection of work to be completed up to project submission** |  |  |
| 1. This item sought to confirm the work that was to be done in each week leading up to the project submission date. 2. Work projection. Team Member, Shi charted out key milestones of the Project from week 7 (current week) to week 11 (presentation and exam week) with Team Member, Chen and Team Member, Lee providing inputs on the deliverable’s details. The projection was tabulated in table 1.   *Table 1: Projection of work*   |  |  |  | | --- | --- | --- | | Week | Deliverables | Work | | 7 | Proposal | Minutes of 1st meeting  Proposal | | 8 |  | Data Cleaning | | 9 |  | Apply analysis on dataset using JMP | | 10 | Report  Poster | Complete Report  Complete Poster | | 11 | Presentation | Rehearse for Presentation |  1. Recurring meetings. The team agreed to meet every Saturday to discuss and level up each other on the project progress. The next meeting date would be on 8 Oct 22. |  |  |
| **Item 4 – Immediate allocaiton of work on proposal** |  |  |
| 1. This item sought to allocate the topics for each member to work on for the proposal. 2. The team discussed and agreed on the follow allocation.   *Table 2: Allocation of work for proposal*   |  |  | | --- | --- | | Section | To be done by | | Project Definition | Bryant | | Data | Yiman | | Methodology | All | | Analytical Tookkit | Chee Liang | | Scope of Work | Chee Liang (Bryant to add on) |  1. The proposal was projected to be drafted by 7 Oct 22 for all members to review. Final adjustments would be made over the following 2 days before submission on 9 Oct 22. | All |  |
| **CONCLUSION** |  |  |
| 1. With no other matters, the meeting concluded at 1615hrs. 2. The next meeting would convene on 8 Oct 22. |  |  |

Recorded by : Shi Chee Liang, Team Member

Distribution:

Shi Chee Liang, Team Member

Bryant Philippe Lee, Team Member

Chen Yiman, Team Member