**AP Supervisor Meeting Minutes – 07 March 2016:**

|  |  |
| --- | --- |
| Date: | 07 March 2016 |
| Time: | 10:30 a.m. |
| Venue: | SMU SIS Meeting Room 4-1 |
| Attendees: | Denise Quek Si Ying |
|  | Tan Wei Song |
|  | Prof. Kam Tin Seong |
| Agenda: | 1. Discuss interim report feedback |
|  |  |
|  |  |

1. **Review of meeting’s discussion**

|  |  |  |  |
| --- | --- | --- | --- |
| **S/N** | **Issues Discussed** | **Action By** | **Due Date** |
| 1. | **Administrative Issues*** Submit original project raw data to Prof. Kam so that he can do a comparison with the cleaned data
* Prof. Kam signed NDA hard copy. To scan the hard copy and forward it to sponsor.
* To put group number (i.e. Group 18) on the cover page of any submissions.
* To change Title of project from “Optimizing…” to one that is more reflective of project – project is about exploring factors that contribute to affecting the purchase/buying patterns of Personal Care
 | Denise & Wei Song | 8 March 2016 |
| 2. | **Review Past Research Papers*** Prof. Kam advised us to look up on past research papers to learn how to improve on our interim report and subsequently, our final report.
	+ E.g. Yeung & Yee Logistic Regression – Logistic Regression on Advancement of Predicting Consumer Purchase Propensity
	+ Look at their introduction, purchase context (how they lead to their hypothesis), data, etc.
* However, research papers are just for reference. Cannot copy wholesale as it will not completely fit into our scenario. E.g. in our case, we don’t have a hypothesis to start with. Therefore, we need to exploratory modeling first and then eliminate insignificant variables from the model.
 | All to note | - |
| 3. | **Overview, Motivation, Objectives*** Take on analytical perspective instead of sponsor’s perspective
* Remove Introduction & Project Background part since it is too high level and contains a lot of business slogans 🡪 Does not explain in our perspective why we take on this project which is to build an effective explanatory model
* Motivation: Change to: Want to build an exploratory model, but problem is that there are too many variables with small records, so we need to figure out how to be able to build an effective explanatory model to explain the factors that affect people in buying certain products
	+ Need to link business issue (Objectives part) with analytical issues
* Include the following to report:
	+ Sampling method – does the sampling require at least xxx number of people that fit into certain demographics, etc.
	+ Sampling frame – What is the frame of consideration? E.g. focus on city areas?
	+ How survey is conducted
	+ End response rate for survey
* Objectives:
	+ The question “who are our target consumers?” is too vague, we need to change it to something more specific, e.g. “Social demographics and economic profile of consumers”
	+ E.g. what are the modes that our target consumers use?
	+ E.g. behavior of consumers?
	+ E.g. to build an explanatory model that allows us to quantify the behavior with the outcome (i.e. purchase pattern)
 | Denise & Wei Song | 14 March 2016 |
| 4. | **Data Preparation*** Currently, it is too confusing for an average reader to comprehend.
* Need to remove the bullet points and group and display them in proper sub-headings (e.g. problems faced, how we overcome it, irrelevant data, outliers) and paragraphs
* Include screenshots to make it clearer as well
* Remove “exclude variables that are not statistically significant…” paragraph 🡪 it should be under methodology instead of data preparation.
* Outliers: cannot just state that how we recode it, but need to provide more explanation on what causes us to recode it? E.g. due to coding error/ extreme values/ statistical outliers but acceptable, etc.
 | Wei Song | 14 March 2016 |
| 5. | **Methodology*** Include explanations on what we are doing in all those stages listed in the process diagram
* We cannot say that we choose a certain model at the very beginning of the project. Must go through the predictive modeling process (i.e. evaluate models already then decide which model to use)
* For recursive partition, include: what are the methods we are going to use? How many trees to build? How many nodes to use?
* For logistic regression, include: Since there is a high volume of variables, are we going to use step-wise? Forward? Backward?
* Model evaluation: What are the criteria to use to evaluate the robustness of our model? Evaluate through misclassification rate? ROC? Confusion Matrix?
	+ Logistic Regression: look at parameter estimates, maximum likelihood
	+ Recursive Partition: look at log-worth
 | Denise | 14 March 2016 |

1. **Next meeting’s agenda**

|  |  |  |
| --- | --- | --- |
| **S/N** | **Issues Discussed** | **Action By** |
| 1. | Next supervisor meeting will be held next week. Date to be confirmed once prof’s schedule has been made available to us. By then, prepare for next week’s meeting based on the agenda stated below.Agenda:* Revised Interim Report
* Wrap up objectives 1 & 2
* Way forward for objectives 3-5
 | Denise, Wei Song |

The meeting was adjourned at 11:30pm. These minutes will be circulated and adopted if there are no amendments reported within the next three days.