Meeting minutes
- Team SkyTrek

Date (dd.mm.yyyy)	23.02.2016
Attendee(s)	Jedaiah, Aseem, Prof Kam
Absentee(s)	Huy

Contents

Item	Discussion point	Remark
1	- Proposal Feedback Clarification	**Feedback:** Only provide the gantt chart without any clarification **Clarification:** Scope of work should be included. To contain the high level requirement, followed by the detailed breakdown of activities required to achieved to achieve the high level requirement Each of these activities would then go into the Gantt Chart Jedaiah to add in scope of work (classification)
2.	- Mid-term report clarifications	**Question:** Should the mid term report focus on demonstrating updates to previous report or be an updated version of the report? **Answer:** Updates. Talk about - how much we have done. - finished data compilations - completed eda? - shared findings with client
3		**Question:** What kind of information is required under the section requiring reasons for selection of a tool **Answer:** - Cost - Skill set requirements - Algorithm implementation differences - Objective evaluation of tools against one another

4		**Question:** Are we going to be assessed on project management adherence? **Answer:** Minutes are important. Used to signal to prof kam how many iterations the team has gone through. Links to the mintutes should be posted on the wiki.
5		**Question:** Will the report be shared online? Concerns over the nda **Answer:** To provide Prof Kam with the NDA so that it wont be shared

Actions

Task	Assigned to	Status
Get Skyscanner NDA to Prof Kam	Aseem	Created
Wiki update for week 8	Aseem & Jedaiah	Ongoing
Manual classification of content types based on heavy weightage terms generated from k means model	Jedaiah	Ongoing
Explore details of incorporating categorical and numerical data into regression model	Huy	Ongoing
Book consultation slot with Prof Kam for week 8 Tuesday this Sunday 1. To clarify with Prof Kam on issues mentioned in proposal feedback How should we account for the NDA since the final conference paper would eventually be published online?	Jedaiah	Done
Complete wiki updates for week 7 1. Minutes clean up and upload Updated Gantt chart	Huy	Done
Complete data pull of 7 datasets and get them verified by Aseem Remarks: Huy will be out of town for week 8 and must ensure that pulled dataset is accurate and ready for analysis in week 8	Huy	Done
Pull Google Analytics Data	Aseem	Done
De-aggregate shares count for each URL	Jedaiah & Huy	Done
Build regression model based on current dataset	Aseem	Done
Come up with list of possible title categories to clear with sponsor	All	Done
Clarify with Prof Kam: Is there a way to dynamically assign new articles to the relevant clusters based on the trained model?	All	Done

Characteristics to scrape from article:
    1.   Number of words (remove stopwords)
    2.   Number of links
    3.   Images
    4.   Videos
    5.   Number of shares of article (by Jed)
	Huy	Done

Retrieve URL share count	Jedaiah	Done
Put topic modelling and google trends research and implementation into Gantt chart timeline	Jedaiah	Done

Input meeting with client in gantt chart for:
    1.   25th Jan
    2.   2nd Feb
    3.   11th Feb
	Jedaiah	Done

Create flow chart documenting process how skyscanner team plans for the next quarter's content postings. To send it to client for verification	Aseem	Done
Test Google API public holiday pulling	Huy	Done
Create a mockup for the following visualization: o To plot Unique page views views (count) vs Week of article publishing o Clicking this particular bar of interest (week has high count) would drill into the next bar chart	Jedaiah	Done

(Number of unique page views, vs identifying URL) o Further clicking on the URL would reveal characteristics of the article (e.g. number of images, videos, etc, bounce/exit rate, etc)		
Create a mockup for the following visualization: o Identify key features of popular posts (MLR against unique page views)	Huy	Done
Create a mockup for the following visualization: o Where do they come from (On the whole country news page level, where do the traffic come from) o How long do they spend at the page (average time on page) o Do they visit other pages? (bounce rate vs exit rate) o How effective are the recommended pages? (If GA does not track the unique page views, then this analysis is not going to be implemented) o [Just to prove a logical point] Correlation test between number of shares vs unique page views	Jedaiah	Done
Deal with proposal segments: 1. Sponsor and Background Information 2. Motivation 3. Objectives 4. Data 5. Scope of Work	Aseem	Done
Questions for Prof Kam: 1. Ask Kam about requirement to post sample date. Concern: data confidentiality 2. What is required for scope of work? (different from methodology)	All	Done

The minutes were prepared by: Jedaiah Tan