COMP257 Project
Analysis of Macquarie University OneHelp Tickets

Kasun Silva, 44810903
Kyle Long, 44989512

Summary

OneHelp is the University wide online request system. OneHelp allows use of the internet to log issues and makes requests, which a person can monitor to completion through automatic email notification and status updates. OneHelp triages and tracks queries, directing them to the correct department. This project will analyze the current OneHelp ticket data, visually displaying findings related to the service teams performance. The project will also define a predictive model that will be able to determine difficulty of a task with a corresponding estimated time of completion.

Project Goals

This project aims to analyze service ticket data from Macquarie universities OneHelp system with the goal of identifying key statistics that can allow for better resource allocation and performance by the service team. Features of interest include:

- Distribution of work by group (HR, IT service desk, Property)
- Distribution of tasks by category (iLearn, Hardware, iPrint, Invoices)
- Efficiency of service groups based on completion times
- Visual overlay of request type over map by location
- A predictive model to determine the difficulty of a task and an expected finish time

The evaluation of request types with location overlaid on a map will visually display resource demands currently and provide insight for future planning of the university. Utilising requests based over time we can help plan better resource allocation and scheduling of the service team. The presentation of these features is of critical importance to the project, it would be ideal at the completion of the project that the discoveries made could be presented to the service team to increase proficiency when dealing with service tasks.

Data

Data obtained directly from official service-now website https://mq.service-now.com (Permission to use this data authorised by FBE IT service Team on the condition that there is no personal identifiable data used). The data was filtered using these conditions and then exported and downloaded directly as a .csv file. The data is in good condition for analysis where most of the columns are either numbers, dates, categories or booleans. There is however the columns "short_desctription" and "location" which will require further processing of the strings to obtain usable information.

	A	B	C	D	E	F	G	H
1	Number	Created	Updated	Short description	State	Assignment group	Active	Closed
2	TASK0475834	06/10/2018 12:17	06/10/2018 12:20	Other ICT Request: Please grant access to SMUAT environment	Closed Complete	IT Service Desk	FALSE	06/10/2018 12:20
3	TASK0475833	06/10/2018 12:12	06/10/2018 12:16	Office365 Mail: FW: Optus Macquarie University Cyber Security Hub Scholarship Offer	Open	IT Service Desk	TRUE	
4	TASK0475832	06/10/2018 12:08	06/10/2018 12:08	OneID: OneID password reset request	Closed Complete	IT Service Desk	FALSE	06/10/2018 12:08
5	TASK0475831	06/10/2018 12:05	06/10/2018 12:05	OneID: OneID password reset request	Closed Complete	IT Service Desk	FALSE	06/10/2018 12:05
6	TASK0475830	06/10/2018 12:04	06/10/2018 12:04	eStudent: request to reset eStudent password	Closed Complete	IT Service Desk	FALSE	06/10/2018 12:04
7	TASK0475829	06/10/2018 11:50	06/10/2018 11:50	eStudent: e-student login	Open	IT Service Desk	TRUE	
8	TASK0475828	06/10/2018 11:48	06/10/2018 11:48	OneID: Can't login to Library multisearch	Closed Complete	IT Service Desk	FALSE	06/10/2018 11:48
9	TASK0475827	06/10/2018 11:44	06/10/2018 11:45	eStudent: Old Password Forgotten	Closed Complete	IT Service Desk	FALSE	06/10/2018 11:45
10	TASK0475826	06/10/2018 11:38	06/10/2018 11:38	eStudent: request to reset eStudent password	Closed Complete	IT Service Desk	FALSE	06/10/2018 11:38

Techniques

Techniques intended to use for predictive model:

➔ Regression: linear/multiple/logistic/k-nearest neighbours
   ◆ Use regression techniques to determine correlation between request types, departments and assignment groups.

- → Unsupervised Learning: k-means clustering
    - ◆ Determine resource allocation based on clustering found in the dataset.
- → Time Series Modelling
    - ◆ Inspecting requests over time to determine if there are trends.

Plan

1) Collect all the data and analyze:

   There are approximately 1.7M tickets in the system and it is only possible to download in 30k chunks. Once all the data has been collected a preliminary analysis of the structure of the data will be done where categories and groups will be identified. The task will be considered complete when there is an identified list relevant categories and groups identified (to be completed 14/10/18).

   We will determine success of categorisation and grouping based on the percentage of tickets that have not be normalised to the whole dataset.

2) Build predictive model:

   This looks to be a difficult task as semantic analysis on "short_description" will be required. The semantics information will then be used to populate new columns that are usable when training models. Multiple models will be used in the project and only the best will be selected as the final model. This task will be considered complete when it is possible to train multiple models on the refined data and the most accurate model is selected (to be complete 21/10/18).