Lark Technologies

Data Analysis Take-Home Assignment

January 2021

1 Introduction

Users in the Lark Diabetes Prevention Program (DPP) are able to interact with the Lark app in many ways. They may have conversations with the chatbot, log any meals they eat, log physical activities, and record their weights. The app also passively records how many minutes users are physically active, and sends notifications to the users periodically.

As DPP users interact with the app they are also exposed to content we call "missions". A mission is a set of conversations about a single topic (e.g. "Get Active", "Eat Well", "Manage Stress") which requires at least 5 completed conversations with the app before a user can move to the next mission. The order of the missions is the same for all users. A user cannot see more than one mission related conversation per day, so one cannot complete a mission in fewer than 5 days. Note that not all in-app conversations are about the user's current mission.

We are interested in understanding how our users interact with the app, why and when engagement drops off, what impact the mission feature has on engagement, and any other related insights this data might provide.

The data is stored in two MySQL tables:

1. user_mission: Information about when a user started and finished a mission.
2. agg_info: User-daily engagement metrics. These engagement metrics are briefly described above.

Please submit your documented code (if using a Jupyter Notebook, also include as a pdf) and a summary of your findings. Make sure to clearly state any assumptions you may have made during your analysis.