Title: A Trajectory-Based Approach to Predict Churn in Online Health Communities

Online Health Communities (OHCs) have positively disrupted the modern global healthcare system as patients and caregivers are interacting online with similar peers to improve quality of their life. Social support is the pillar of OHCs and, hence, analyzing the different types of social support activities contributes to a better understanding and prediction of future user engagement in OHCs.

This study used data from a popular OHC, called Breastcancer.org, to first classify user posts in the community into the different categories of social support using Word2Vec for language processing and six different classifiers were explored, resulting in the conclusion that Random Forest was the best approach for classification of the user posts. This exercise provided useful inputs about the nature of web interactions of users in the community. Furthermore, it also helped identify the most common type of social support activity among users.

Thereafter, a trajectory-based Bayesian model was proposed and implemented to predict user churn (attrition) from the OHC. Comparison of the proposed trajectory-based method with three other benchmark methods established that user trajectories, which represent the month-to-month change in the type of social support activity of users, are effective pointers for user churn from the community.

The results and findings from this study could help OHC administrators better understand the needs of users in the community and take necessary steps to improve user retention and community management.