

A Data-Driven Approach to Understanding Happiness

Akari Asai¹

¹Univ. of Tokyo

akari-asai@g.ecc.u-tokyo.ac.jp

Vivian Li²

Daniela Stepanov²

Wang-Chiew Tan²

² Recruit Institute of Technology

{vivian, daniela, wangchiew}@recruit.ai

1 Introduction

The science of happiness is an area of positive psychology that studies the factors that sustain people’s happiness over time (Seligman, 2011; Fredrickson, 2009; Lyubomirsky, 2008). One of the interesting findings of the field (Diener et al., 1999) is that while 50% of our happiness is genetically determined, and only 10% of it is determined by our life circumstances (e.g., finances, job, material belongings), 40% of our happiness is determined by behaviors that are under our control. Examples of such behaviors include investing in long-term personal relationships, bonding with loved ones, doing meaningful work and caring for one’s body and mind. Consequently, positive psychologists have focused on devising methods to steer people towards those behaviors.

There has been recent interest to develop technology that helps users incorporate the findings of the science of happiness into their daily lives. Current applications fall into two categories. In the first, the application tries to suggest relevant content to the user based on their answers to a predefined set of questions (Killingsworth, 2017; Happify, 2017). In the second, users can log their emotions in a journaling-style environment but that content is available mostly for their own reflection (Bliss, 2017; DayOne, 2017).

At the Recruit Institute of Technology, our goal is to broaden the data we collect about people’s happiness to include text that describes their happy moments. Understanding such text is crucial for an interactive assistant to be able to provide insights that are tailored to the user’s specific life circumstances. In what follows we describe Joybot v1.0, the first version of our happiness app that allows a user to record their happy moments via text, voice, or photographs on one’s mobile device. Joybot interprets these happy moments au-

tomatically and develops, as happy moments are entered, an understanding of the key factors that make the user happy and how. We then describe the technical NLP challenges that Joybot raises, and a crowdsourced corpus of happy moments that we collected in order to advance the research on NLP for happiness.

2 Joybot: a happiness app

Joybot is an interactive journaling application that aims to provide useful advice to the user by analyzing the user’s journal entries. Users interact with Joybot via two main interfaces: visualization and recommendations. Joybot’s visualizations are motivated by the need to provide constructive feedback to users regarding happy moments that are entered. This is done by categorizing the happy moments into “happiness” categories, and providing users with more insights into their activities and how they change over time. We chose a set of 7 categories based on studying the psychology literature and inspecting thousands of happy moments contributed by crowd workers. The categories are: exercise, affection, bonding, nature, leisure, enjoy the moment, and achievement. We developed a multi-label classifier using Logistic Regression as our binary classifier for each category.

The second interface Joybot aims to support is a recommendation system based on identifying patterns in past happy moments. For example, if the user frequently reports being happy when spending quality time with his/her daughter, Joybot may recommend doing more of the same. Building such an interface raises several NLP challenges: we need to understand the activities that make the user happy even if they are reported in differing surface forms and at different levels of granularity. In our example, the user will not typically report

mom we need to extract the direct object of the root (dinner) along with the root verb (had) in order to fully capture the activity “had dinner”. In general, there could be multiple verb phrases in the happy moment, and deciding which one of them conveys the main activity can be subtle.

We have developed a model based on Conditional Random Fields (CRF) and trained it on labeled data to extract key activity. Our labeled data consists of about 500 happy moments, each manually annotated with the key activity. We plan to improve the extraction of key activities by considering semantic role labeling using Propbank (Palmer et al., 2010) or other types of frames. A better understanding of happy moments through frames will help guide the extraction of key activities. In addition, we also plan to investigate more directed extraction methods that leverages prior knowledge of the categories that a happy moment belongs to. We have developed an algorithm that classifies a happy moment into its most appropriate category (e.g., Affection, or Nature, or Bonding etc.) For example, if a happy moment belongs to Affection, then we know it is likely that there are people or pets involved and we are hence able to make a directed effort to extract those content. Finally, with a good understanding of what are the key activities of a simple happy moment, we can then understand what are the main activities of more complex happy moments.

Second, a more conceptual challenge is that, even for the very simple happy moment *I had dinner with my husband*, the potential sources of happiness could be as diverse as eating dinner, being with the husband, or even a factor that is not mentioned in the text, but could be inferred from others, such as having a date night without the children. Furthermore, in order to fully understand the happy moment we need to extract the participants mentioned in the moment (and decide whether they played a key role in the happiness or not). Finally, many moments include text that is superfluous to the cause of happiness. For example, in the moment *my daughter woke up this morning, came outside and helped us with the yard work this morning*, most of the text surrounding the main activity is superfluous. We posit that a better understanding of main activities will help us to distinguish important clauses from the superfluous ones.

In conclusion, we do not expect that we can de-

velop NLP techniques that completely understand the causes of happiness in every scenario. However, we hope to develop technology that is useful enough to provide users valuable feedback. From a business perspective, understanding the causes of happiness related to products and services from comments on social media by advertisers or third parties can also help business owners refine their products and services to increase the satisfaction of their customers.

References

- Tim Althoff, Kevin Clark, and Jure Leskovec. 2016. Large-scale Analysis of Counseling Conversations - An Application of Natural Language Processing to Mental Health. *TACL*.
- Bliss. 2017. Gratitude journal - bliss. <http://bliss31.com/>.
- DayOne. 2017. Day one - a simple and elegant journal for iphone. <http://dayoneapp.com/>.
- Diener, Suh, Lucas, and Smith. 1999. Subjective well-being: Three decades of progress. *Psychological Bulletin* pages 276–302.
- Barbara Fredrickson. 2009. *Positivity: Top-Notch Research Reveals the Upward Spiral That Will Change Your Life*. Harmony.
- Happify. 2017. Happify: Science-based happiness games activities. www.happify.com/.
- Matt Killingsworth. 2017. Track your happiness. <https://www.trackyourhappiness.org>.
- Huijie Lin, Jia Jia, Liqiang Nie, Guangyao Shen, and Tat-Seng Chua. 2016. What does social media say about your stress? In *IJCAI*.
- Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies* 5(1):1–167.
- Sonja Lyubomirsky. 2008. *The How of Happiness: A New Approach to Getting the Life You Want*. Penguin Books; Reprint edition (December 30, 2008).
- Rada Mihalcea and Hugo Liu. 2006. *A corpus-based approach to finding happiness*. In *Proc. AAAI Spring Symposium and Computational Approaches to Weblogs*. page 6 pages. <http://www.cse.unt.edu/rada/papers.html>.
- Martha Palmer, Daniel Gildea, and Nianwen Xue. 2010. *Semantic Role Labeling*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers.
- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval* 2(12):1–135.
- Martin E. P. Seligman. 2011. *Flourish: A New Understanding of Happiness, Well-being - and how to Achieve Them*. Nicholas Brealey.
- Martin E. P. Seligman, Tracy A Steen, Nansook Park, and Christopher Peterson. 2005. Positive psychology progress: empirical validation of interventions. *American psychologist* 60(5):410.