Proposal

The Ebbinghaus model, also termed "the forgetting curve", posits that memory decays exponentially over time (Ebbinghaus, 1885):

$$p = 2^{\frac{-\Delta}{h}} \tag{1}$$

Settles and Meeder (2016) proposed a half-life regression machine learning model based on the Ebbinghaus model to improve Duolingo daily student engagement by 12% in an operational user study. We aim to use replicate the approaches described in this study to analyze a dataset of 13 million Duolingo student practice sessions, ultimately being able to predict how well a student will remember or not remember words over time. Thus, our model will have the ability to predict an individual's memory decay to optimize their frequency of practice. The key behaviors that will be analyzed are native language, recall rate, time between practice sessions, and word difficulty. Extending this research, we will also aim to test out other machine learning models in accordance to the Final Project guidelines.

References

Ebbinghaus H. Memory: a contribution to experimental psychology. Ann Neurosci. 2013 Oct;20(4):155-6. doi: 10.5214/ans.0972.7531.200408. PMID: 25206041; PMCID: PMC4117135.

A Trainable Spaced Repetition Model for Language Learning(https://aclanthology.org/P16-1174) (Settles & Meeder, ACL 2016)