Spaced Repetition Based Adaptive E-Learning Framework

Amit Kharwal[1], Neelay Umrotkar[2], Varun Godambe[3],
Uttam Kolekar[4], Vishal Badgujar[5]

Department of Information Technology
A.P. Shah Institute of Technology
Thane (M.H.), India
([1]amitkharwal,[2]neelayumrotkar,[3]varungodambe,
[4]udkolekar,[5]vsbadgujar)@apsit.edu.in

Abstract. With global education moving towards online education, we see a lot of emerging techniques to help facilitate paradigm shift. Memorization being a fundamental part of the process of learning, we focus on making memorizing an efficient task for students. We plan to make use of 2 emerging methodologies that help us memorize and retain information better in the long term. These are: Spaced Repetition: this technique helps in retaining information efficiently. Adaptive Learning: This will help us make the process of learning personalized, so that the user is always challenged just enough to enable growth. To help users stay consistent, we plan to add game-like elements to the system to help facilitate regular use which will lead to better results.

Keywords: Spaced Repetition · Adaptive Learning · Memorization · Re-Collection · Flashcards · Machine Learning.

1 Introduction

The Indian education system inherently places a lot of importance on memorization. Tests are designed to be subjective that require long descriptive answers which in turn need the user to either understand the topic to the core or memorize sets of information which they can string together during the test. This pattern is visible in throughout a student's learning career. Often one of the things that happen most often is that students rote-memorize information haphazardly and fail to reproduce it in the tests, and even if they do reproduce it properly, they do not recall it later on after the exams. This is a sub-optimal and incorrect way to go about testing and memorization/learning in general. We intend to develop a system that helps students memorize and retain information in a systematic and efficient way. In order to do this, we decided to use techniques like Spaced Repetition, Adaptive Learning and in conjecture with these, we plan to add game like elements to keep the users' interest so as to not miss out on practice. Below, we've explained how each of these techniques mentioned

above will play a part in our project.

Spaced repetition is a technique for efficient memorization which uses repeated reviews of content following a schedule determined by the spaced repetition algorithm to improve long term retention. We review information in the form of cards of information, at gradually increasing intervals. For example, if you answer a card correctly, the algorithm will repeat that card less frequently, whereas if you get the answer wrong, it'll be asked more frequently. The reason this technique works better is that the revision is done at or around the exact time when the user is expected to forget said information. Spaced Repetition is significantly more effective than rote learning.[11] .

Adaptive learning instead of a static and rigid learning system which forces everyone to follow the same learning pattern, you're given a personalized experience. If the learner takes a test, the algorithm can adapt the training according to the answers given. Doing badly on certain questions can trigger the system to push relevant content to the learner, or the overall performance could determine which learning path they take. Adaptive learning refers to measuring and responding to each activity while a student progresses in the material assigned by their instructor. When combined with data from student interactions, it is able to present the best content (right type, right amount, right level) for that moment in time to keep the student engaged. Flexible adaptive learning can help preserve the content regardless if you're teaching on campus, online or a mixture of both. Rooted in several learning science principles like spacing, chunking, and interleaving, create mini- cycles of questions that create smaller, easier to absorb chunks of content. If students continue to struggle with a concept, they are more likely to encounter repeating questions. This method can help instructors understand their students better by understanding the data provided by the quiz results.[6]

Learning Curve: Two of the most common methods we see in the market are:
1) Pimsleur Method: Here you are asked to retain information learned after a predefined interval, without taking into account your performance in the previous review. This method is extremely restrictive and one-dimensional.
2) Leitner Method: Here, a few different intervals in terms of boxes, are predefined. Let's say we have the intervals of 1, 2, 4, 8, and 16 days. Every new piece of information starts in the 1 -Day Interval Box. If the user remembers the information after 1 day, the information gets put in the 2 - Day Interval Box. If the user remembers the information after 2 days, it gets promoted to the 4-Day Interval Box, so on and so forth. If at any moment, the user fails to recall the information, then the piece of information gets shifted to the previous interval box. This method was a massive upgrade over the Pimsleur Method. But it can still be made even better by using another method.

2 Objectives

1. To design an adaptive algorithm that takes user performance into account in order to adapt the next question's levels.
2. To design a space repetition algorithm that takes user's specific quiz performance into account.
3. To integrate Adaptive and Spaced repetition algorithms for the proposed system.
4. To add game design elements to make the layout appealing.
5. To create a cross-platform application interface.

3 Literature Review

In literature [4], a centralized system is proposed which will act as a portal for both students and instructors where the instructors can curate content according to their student's needs and the students can access this content. The performance of students on using this content will then be available to the instructors to then modify their offered content. The constant cycle of content modification according to the needs of the student and analysis of the student's performance greatly interest us. Keeping content relevant will be an important part of the memorization process and so will the analysis.

In literature [5], The literature ponders over how our traditional learning systems could implement a cloud-computing based e-learning system. It discusses the advantages of cloud-computing based architecture such as high availability which is definitely a value that we hold important in our own implementation.

In literature [6], the author focuses a significant chunk of itself on explaining the deference between personalized learning and adaptive learning, where, adaptive learning focuses on individual differences, individual performance and adaptive adjustment whereas personalized learning focuses on all of these but with the addition of personal needs. This here, makes a big difference. As our system focuses more on memorization and learning rather than teaching, one of the things we can adopt from this research is a feedback system which helps specialists and teachers cater to the student's needs better.

In literature [7], our takeaway is to increase the amount of content available for the user thereby reducing the chances of repetition in the same session while also removing the chances of the user cramming all the content of the application in one go, thereby maintaining the integrity of the spaced repetition-based system.

In literature [8], the author focuses on the effects of adaptive learning in the domain of language learning and teaching. In adaptive learning, learner's interaction with previous content determines subsequent content. Hence a sort of personalization is seen. Digital flashcards are used to incorporate automated spaced repetitions of the targets learning item. Since the system is individualized, it can adapt relatively easily in response to small amounts of data. Stated to be a 'hot concept' that is 'poised to reshape education' in Webley magazine. As we

intend to make the application based on digital flashcards, we can use the idea of creating algorithms to determine the order, frequency and difficulty of the next learning item based upon their previous exposure to it.

In literature [9], the author provides and details an adaptive learning platform which was implemented from scratch and used for undergraduate courses related to digital systems and computers. Students differ in their learning capabilities and it will be difficult to provide a tailored education, especially in large classes. Different technology solutions have been proposed to deal with this challenge. Rule based approaches use pre-determined branching architectures based on his response to questions. If a question is answered correctly, the system could issue a more challenging question which is pre-determined. The system can issue an easier question or hints or other assistance if the question is answered incorrectly. Algorithmic mechanisms use mathematical models for inferring student's ability level perform real- time content adaptation. This paper helped us reinforce that adaptive learning does indeed help students memorize and learn better.

In literature [11], the author's model captures what the student has learned once a module is completed, and estimates how well he can recall this knowledge at any given time. Spaced repetition is a key component of the student model. Over time, the strength of a skill will decay in the student's long-term memory, and this model helps the student manage his practice schedule. The spacing effect is people remembering things more effectively if they use spaced repetition as op- posed to cramming. We are given a brief exposure to primitive spaced repetition methods such as the Pimsleur Method and the Leitner Method. The main re- search that happens in the paper is related to the Ebbinghaus Forgetting Curve[10] and how the researchers tweaked the original curve to make a model that more performs better when calculating the knowledge half -life. We will be using the Ebbinghaus Curve model to has our Half-life calculations on.

In literature [12], the author provides us with a primitive idea of how spaced repetition helps students perform better and how gamification helps to have students engaged during the process of learning and how it also helps them return to the process regularly which in turn also helps us understand just how effective spaced repetition is. We do plan to implement game-like features in our project to showcase similar results of engaging and retaining students to the learning process. We would also, as is regularly done with any spaced repetition program, expose students to maintenance and acquisition cards.

4 Existing System Architecture

Anki Spaced Repetition Tool:[2] SM-2 stands for Super Memo 2[1], which is a spaced repetition algorithm. It was developed my P.A Wozniak. It has multiple versions right from SM-0 to SM-18. Despite these versions, SM-2 remains the most popular. In its calculations, it takes into account factors such as the "repetition number," "easiness factor" and "interpretation interval." Here, repetition number refers to how many times a certain element has been repeated. Easiness factor refers to how quickly the interval between each repetition increases. Inter-

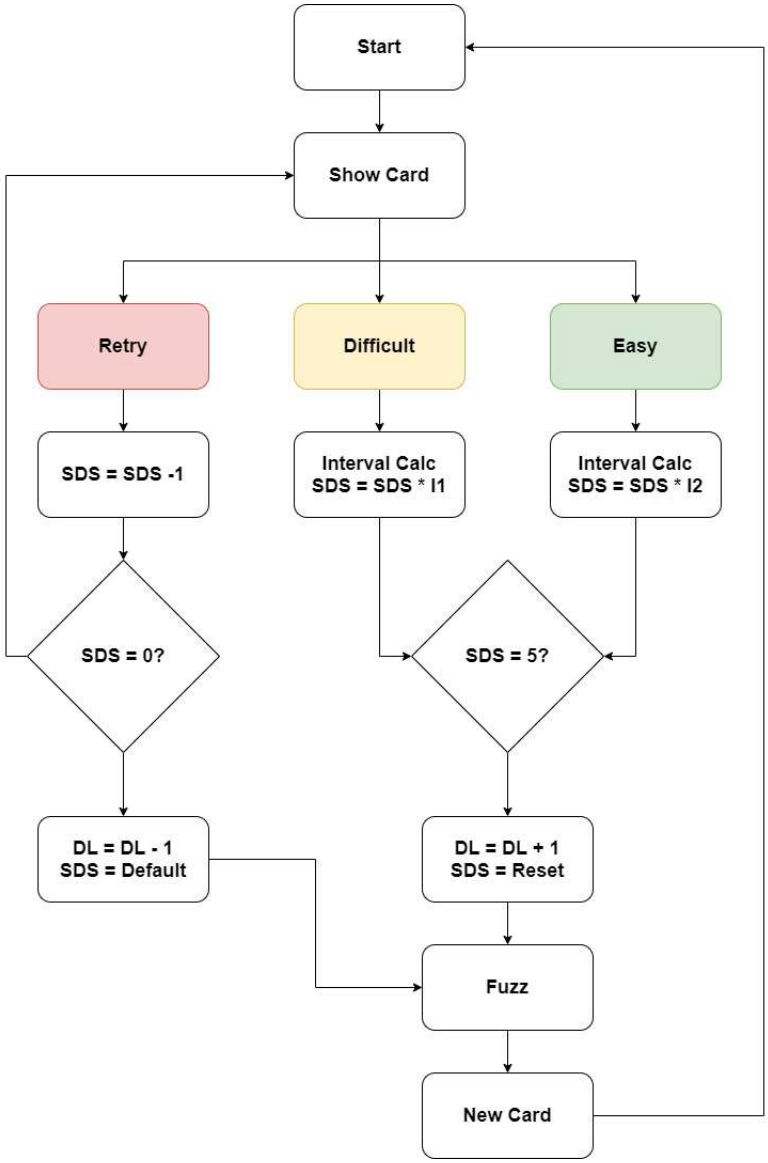

Fig. 1. Existing System Architecture.

Legend	Meaning
E	Ease coefficient
I	Interval
EB	Ease Bonus Coefficient
Fuzz	Small Randomized Addition to Interval to Prevent Repetition of Card Sequence

Fig. 2. Legend for Existing System Architecture.

repetition interval refers to the amount of time between repetitions.
In Anki, they use a tweaked version of the SM-2 spaced repetition algorithm. Once the flashcard is showed on the screen, the user is given 4 options. If the user selects the "Repeat" option, the card is repeated again after the deck is done, and the Ease is decreased by 20 percent points. If the user selects the "Hard" option, the Interval is multiplied by 1.2 and the ease is increased by 15 percent points. If the user selects "Good" option, the interval is multiplied by ease, the Interval remains the same. If the user selects the "Easy" option, the Interval is multiplied by Ease * Easy bonus and Ease is increased by 15 percentage points. Here. It is made that if the person is having difficulty with a certain question, it is repeated more often. The minimum time interval is never less than a day.
Although Anki covers the basics of spaced repetition very well, it lack interest building elements. To cover this, we looked at another language learning tool available on devices as well as on the web called "Clozemaster"[3]. Clozemaster is a tool that helps people learn and memorize vocabulary, but their approach to the same is very different. Where Anki uses the same only cards with a very toned-down visual environment, Clozemaster uses a lot of game-like elements to keep the user's interest. Although they don't used a proper spaced repetition system, the visual language that they use is something that can be learnt from. Elements that we found extremely necessary and helpful:

- Colorful Design
- Progress Bars
- Daily Goals
- Personal Stats
- Leader boards

5 Proposed Architecture

In our enhanced Spaced Repetition based Adaptive Learning Framework, we start of with the user being showed a flashcard. The user has been assigned default values of SDS and Interval. And the questions have the value of DL. Once shown the flashcard, the user has 3 options:
Retry: If the user selects this option, first, his SDS is reduced by 1. Next it is checked if the SDS is equal to Zero. If it is, it means that the user hasn't been able to answer 5 questions of the same difficulty. In short, the questions are too difficult for him. In this case, the DL is reduced by one to ease the user and the SDS returns to 1. Now, new cards of the reduced difficulty are produced.
Difficult: Here, if the user selects this option, the interval is calculated using our half-life calculations based on Ebbinghaus' Forgetting Curve equations and the SDS is increased by 0.5.

Easy: Here, if the user selects this option, the Interval is again calculated using the half-life formula and the SDS is increased by 1.
In both Difficult and Easy options, if the SDS is equal to 5, the DL is increased by 1 thus producing new questions that are more challenging in nature to keep

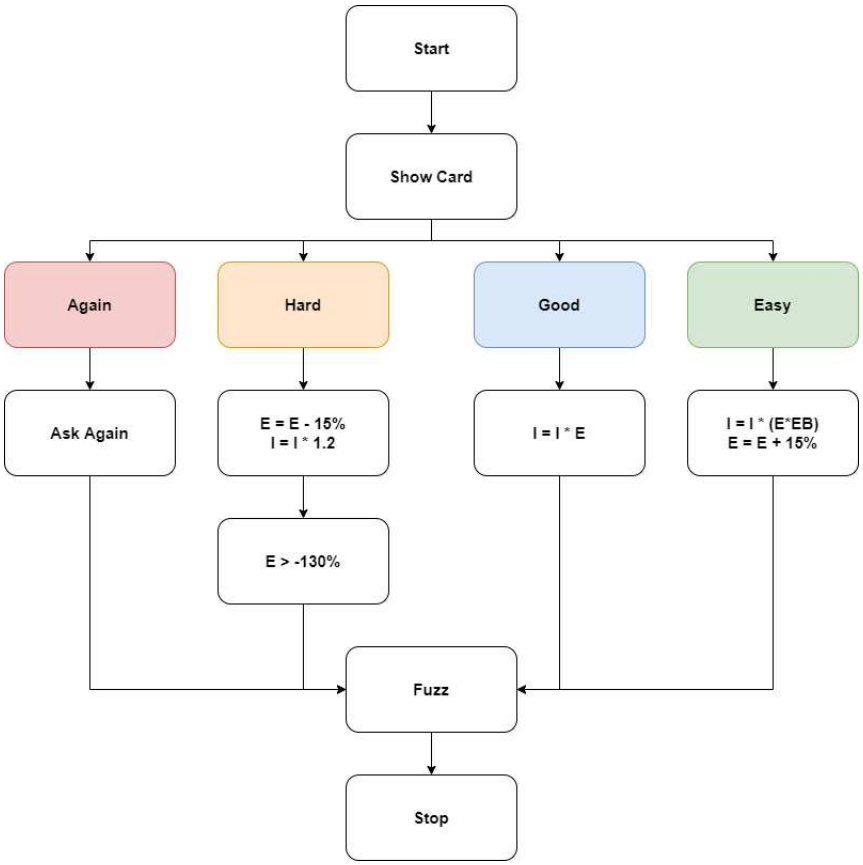

Fig. 3. Proposed Architecture.

the user challenged.In our implementation, we use the half-life calculations to adjust the time intervals according to the needs of every individual.

Another objective we want to achieve is a design that keeps the user engaged. For this we decided to add game-like elements to our interface. Surface level designs that help us retain users. For this, we plan to implement the following elements:
• Leader boards (Local and Global): Here the user will be able to see how they match up to the local and global high scores. This will help them keep motivated.
• Friend's List: Here the user can see their friends. On clicking on their profile, they can see an overview of how their friends perform as compared to themselves.
• Statistical Analysis Page: Here the user can see their performance on different scales across all their subscriptions. Suggestions will also be provided such that the user can improve.
• Progress Bars: Here the user can see a summary of how far they've completed

various topics.

• Achievement Awards: For every milestone, users will be awarded an award. This is a positive reinforcement technique that will help students be consistent. The interval is decided on the basis of how the user responds with their selection of a difficulty with regards to the question they face. Following the Forgetting Curve mapping by Ebbinghaus, we decided to stick to major milestones on the timeline as the interval progression. As shown in the table below, we start with the first repetition after 18 minutes of going over the card for the first time. This is followed by a repetition after 1 day, which is then followed by a repetition after 6 days, 30 days, 90 days and finally 180 days at which point of time, the interval continues to be 180 days. The basic structure follows that of the proposed infrastructure wherein the user is presented with 3 different difficulties and further repetition-interval calculations are done depending on the difficulty chosen.

Retry: if the user selects this option, first we check if the current interval is equal to the minimum interval which is 18 minutes. If it is, the interval will continue to remain 18 minutes till the next repetition. If the repetition is not the minimum 18 minutes, the interval will be demoted to the previous interval. For example, if the current interval is 30 days, it will be demoted to 6 days.

Difficult: If the user selects this option, there will be no changes in the interval. It will remain the same.

Easy: If the user selects this option, first we check if the current interval is equal to the maximum interval which is 180 days. If it is, the interval will continue to remain 180 days. If the repetition is not the maximum, the interval will be incremented to the next interval. For example, if the current interval is 30 days, it will be incremented to 90 days.

6 Conclusion

To summarize, we propose a new system that aids in memorization, of which the 3 core components are Spaced Repetition, Adaptive Learning and Game-Like Elements. The spaced repetition algorithm will feature a more accurate interval calculation which is based on Ebbinghaus' Forgetting Curve which will make sure that the user is asked to recall the target information just as they're about to forget it, leading to better retention in the future. The adaptive learning algorithm will monitor the user's performance. Based on the data that it collects; it will then adjust the difficulty of the questions asked to the user such that the difficulty doesn't deter them away from the learning process but also doesn't make learning too easy. A balance between challenge and reward will be the key to making sure the system works in the intended manner. For this, the system will feature the implementation of game-like elements such as achievements, badges, awards, etc, which promote a sense of achievement and make users want to achieve their individual goals. Apart from this, we will also add features like leaderboards and personal performance statistics that will enable

the users to compete with each other, increasing competition and subsequently also increasing the drive to use the system more often.

References

[1] Supermemo Algorithm. Accessed August 2021. URL: https://www.supermemo.com/en/archives1990-2015/english/ol/sm2.

[2] Anki. *Frequently Asked Questions*. Accessed August 2021. URL: https://faqs.ankiweb.net/what-spaced-repetition-algorithm.html.

[3] Clozemaster. *Frequently Asked Questions*. Accessed August 2021. URL: https://www.clozemaster.com/faq.

[4] Shetu S. F. et al. *Impactful e-learning framework: A new hybrid form of education*. 2021. URL: https://www.sciencedirect.com/science/article/pii/S2666518221000255.

[5] MAH Masud and X Huang. *An e-learning system architecture based on cloud computing*. 2021. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.221.4704&rep=rep1&type=pdf.

[6] Rajesh C. Panicker et al. *Adaptive Learning and Analytics in Engineering Education*. 2018.

[7] Hongchao Peng, Shanshan Ma, and Jonathan Michael Spector. *Personalized adaptive learning: an emerging pedagogical approach enabled by a smart learning environment*. 2019.

[8] Florian Schimanke et al. *Implications of Short-Term Memory Research for the Design of Spaced Repetition Based Mobile Learning Games*. 2015.

[9] Burr Settles and Brendan Meeder. *A Trainable Spaced Repetition Model for Language Learning*. 2016.

[10] J. Whittman. *The Forgetting Curve*. URL: https://www.csustan.edu/sites/default/files/groups/Writing%20Program/forgetting_curve.pdf.

[11] Martin K.-C. Yeh et al. *Using Spaced Repetition and Gamication to Enhance K-12 Student Science Literacy with On-Demand Mobile Short Reads*. 2016.

[12] Bingxue ZHANG et al. *Recognition andApplication of Learner's Cognitive Ability for Adaptive E-learning*. 2020.